Spatial and Rapid Mixing on Lattice Graphs

Thesis submitted in accordance with
the requirements of the University of Liverpool
for the degree of Doctor of Philosophy by

Markus Jalsenius

First Supervisor:	Prof. Leslie Ann Goldberg
Second Supervisor:	Dr. Paul Goldberg

Department of Computer Science
The University of Liverpool

April, 2008

Abstract

In this thesis we consider the anti-ferromagnetic Potts model on lattice graphs. A spin system under this model is characterised by an underlying connected graph and two parameters: q, the number of spins, and $\lambda = \exp(-\beta)$, where β is the "inverse temperature". Two questions of interest are to determine if a spin system has strong spatial mixing and if the Glauber dynamics is rapidly mixing on the configuration space. These two properties are closely related, in particular it is known that strong spatial mixing often implies that the Glauber dynamics is rapidly mixing. Rapidly mixing Glauber dynamics implies that there is a fully-polynomial randomised approximation scheme for the partition function. For graphs in which the distance-d neighbourhood of a vertex grows sub-exponentially in d, strong spatial mixing implies that there is a unique infinite-volume Gibbs distribution.

We use recursively-constructed couplings to derive mixing bounds for any graph and any temperature. The result improves previously known general mixing bounds. Our main objective is to have results which are applicable to the lattices studied in statistical physics. In this thesis we focus on the square lattice ($\mathbb{Z}^2$), the triangular lattice and the kagome lattice. By considering the geometry of the lattice we are able to achieve better mixing bounds. Rather than constructing recursive couplings from a single recurrence, we use a system of recurrences, which is highly dependent on the geometry of the lattice. For the square lattice we give a proof of strong spatial mixing and rapid mixing for $q \geq 6$ and any λ. We also show that mixing occurs for a larger range of λ than was previously known for $q = 3$, 4 and 5. Certain probabilities that are used in the proof are obtained with computer assistance which makes the proof machine assisted.

The anti-ferromagnetic Potts model corresponds to proper colourings when the temperature is zero. By refining the technique of recursively constructing couplings, we provide proofs of mixing for the triangular lattice with $q = 9$ and $\lambda = 0$, and the kagome lattice with $q = 5$ and $\lambda = 0$. This improves previously known mixing bounds. The systems of recurrences we use here are rather large and require a computer to be constructed. This makes the proofs machine assisted. The Glauber dynamics is not necessarily irreducible on the kagome lattice with $q = 5$ and $\lambda = 0$ if we impose a boundary condition. However, we show rapid mixing under the free boundary.

We also study the mixing time of a systematic scan Markov chain for sampling from the uniform distribution on proper 7-colourings of a finite rectangular sub-grid of the square lattice. A systematic scan Markov chain updates finite-size subsets of vertices in a deterministic order. We use a heuristic-based computation in order to establish a rigorous result about the mixing time. The proof is computer assisted and improves previously known mixing bounds for systematic scan on the square lattice.

In memory of my father

Contents

List of Figures

Preface

The results presented in this thesis are on the whole the product of my own work. However, some of the results come from joint work with others. Chapter 3 summarises the contribution by this thesis in terms of theorems that improve previously known results. Theorem 3.1 and 3.2 are due to Goldberg, Jalsenius, Martin and Paterson, and are published in [GJMP06]. Theorem 3.6 is the result from joint work with Pedersen, to which we have both contributed equally.

The idea of showing strong spatial mixing and rapidly mixing Glauber dynamics for lattice graphs by recursively constructing couplings, represented as trees, was introduced by Goldberg, Martin and Paterson in [GMP05]. We exploit their idea in this thesis in order to derive new improved mixing bounds for certain spin systems. We use notation from the very same paper, although we modify it slightly such that it will suit the contents of this thesis. It is clearly stated when ideas are borrowed from others.

Acknowledgements

I would like to express my sincere gratitude and appreciation to my supervisor, Leslie Ann Goldberg for directing me into a stimulating and challenging research area, and for her continuous support and for always being available when I needed her advice. Her knowledge and her logical way of thinking have been of great value for me.

During my Ph.D. studies, I have been a part of two great research groups: the Algorithms and Computational Complexity Research Group in the Department of Computer Science at the University of Warwick, and the Complexity Theory and Algorithmics Group in the Department of Computer Science at the University of Liverpool. I would like to thank both groups and departments for providing an excellent and very pleasant environment in which to conduct research. In particular, I would like to acknowledge Russell Martin and Mike Paterson with whom I had the opportunity to collaborate.

I thank the University of Liverpool and the Department of Computer Science at the University of Warwick for providing the funding underlying my Ph.D. studies.

The time spent in the department would never have been the same without three special officemates: Nick Palmer, Kasper Pedersen (with whom I really enjoyed coauthoring a paper) and Pattarawit Polpinit (also being the ideal housemate).

I would also like to express my gratitude to the teachers I have had since my first day of school. Thanks to them, my journey from seven-year old pupil to Ph.D. student has been possible.

I owe special thanks to all my friends, who I am very fortunate to have.

There are hardly words enough to thank my family, especially my mother and my sister, for their endless support and for always being there for me. I direct a special thought to my father who always inspired and encouraged me to be curious and seek new knowledge.

Notation glossary

The following list of notation is in alphabetical order. For example, $\Gamma^d(X)$ will be sorted as the string "Gamma d X". For each item it is specified on which page it is defined or introduced. The second list covers some definitions.

$1_{\Psi,v}$	An indicator random variable for the event that the colour of v differs in a pair of colourings drawn from Ψ, Page 50
$\mathcal{A}$	A collection of sets of edge-boundary pairs, Page 77
$(A, F, \mathcal{H})$	An element in an $(\mathcal{A}, \mathcal{F})$-set, Page 77
α	In systematic scan: The influence on any vertex, Page 118
$A(\phi)$	The congestion of the flow ϕ, Page 45
$\mathcal{B}$	Colouring of a vertex boundary, Page 8
B	Colouring of an edge boundary, Page 50
$\mathrm{Ball}_d(v)$	The vertices that are at most distance d from $v \in V_\mathcal{G}$, Page 22
$B(E)$	The colouring of edges E induced by B, Page 50
$B(e)$	The colouring of edge e induced by B, Page 50
β	The inverse temperature, Page 11
$(\mathcal{B}_\mathcal{X}, \mathcal{B}'_\mathcal{X})$	The pair of colourings of $\partial R_\mathcal{X}$ in vertex-boundary pair $\mathcal{X}$, Page 50
(B_X, B'_X)	The pair of colourings of $\mathcal{E}R_X$ in edge-boundary pair X, Page 53
$\deg(v)$	The degree of vertex v, Page 52
Δ	The maximum degree of $\mathcal{G}$, Page 8
∂R	The vertex boundary of region R, Page 8
$\mathrm{d_{TV}}(D_1, D_2)$	The total variation distance between D_1 and D_2, Page 14
$d(x, y)$	The vertex-distance between x and y, where x and y can be vertices, edges or sets of vertices, Page 8
$\mathbb{E}[1_{\Psi,v}]$	The probability vertex v differs in two colourings drawn from Ψ, Page 50
$E_\mathcal{G}$	The edge set of $\mathcal{G}$, Page 7
$\mathcal{E}R$	The edge boundary of region R, Page 8
$E(R)$	The set of edges incident to vertices in region R, Page 8
$e_X = (w_X, v_X)$	The distinguished boundary edge in $\mathcal{E}R_X$, Page 52
$\mathcal{F}$	A collection of sets of edge-boundary pairs, Page 77
$\mathcal{G} = (V_\mathcal{G}, E_\mathcal{G})$	The underlying infinite graph, Page 7

$\Gamma^d(X)$	The sum of the likelihoods of level-d edges in T_X, Page 66
$\mathcal{H}$	A set of $(\Delta-1)$-tuples $(A_1,\ldots,A_{\Delta-1})$, such that $A_i \in \mathcal{A}\cup\{\emptyset\}$, Page 77
λ	The temperature parameter ($\lambda \in [0,1]$), Page 13
$\mathcal{M}$	A Markov chain, Page 16
$\mathcal{M}_0$	The Glauber dynamics, Page 37
$\mathcal{M}_d$	The heat-bath Markov chain with radius-d ball updates, Page 37
$\mathrm{mon}_{\sigma,\mathcal{B}}(E)$	The number of monochromatic edges in E under σ and $\mathcal{B}$, Page 13
$\mathrm{mon}_{\sigma,B}(E)$	The number of monochromatic edges in E under σ and B, Page 51
$\mathcal{M}_{\text{scan}}$	The heat-bath systematic scan Markov chain, Page 19
$\mu_{c,c'}(X)$	A value that could upper-bound $\nu(X)$, Page 69
μ_F	A value that upper-bounds $\mu(X)$ over all $X \in F$, Page 78
$\mu(X)$	A value that is upper-bounding $\nu(X)$, Page 69
$\mathbb{N}$	The natural numbers $\{0,1,2,\ldots\}$, Page 20
$\nu(X)$	The value $\mathbb{E}[1_{\Psi_X^{\min},v_X}] = \sum_{c,c'\in Q, c\neq c'} p_X^{\min}(c,c')$, Page 68
n_X	The number of neighbours of v_X that are in R_X, Page 53
Ω	All proper q-colourings of $V_{\mathcal{G}}$, Page 8
ω_i	The weight of colourings with v_X coloured i, ignoring edge e_X, Page 68
$\hat{\omega}_{c,c'}$	The sum $\sum_{i\in Q\setminus\{c,c'\}} \omega_i$, Page 68
Ω^+	All q-colourings of $V_{\mathcal{G}}$, Page 8
$\Omega^{+2}(u)$	The set of pairs $(\sigma_1,\sigma_2) \in \Omega^+\times\Omega^+$ only differing on vertex u, Page 117
Ω_R^+	All q-colourings of region R, Page 8
Ω_R	All proper q-colourings of region R, Page 8
$\Omega_R(\mathcal{B})$	All proper q-colourings of R that agree with $\mathcal{B}$ of ∂R, Page 8
Ω_k^σ	The set of colourings σ' such that $\sigma' = \sigma$ off Θ_k, Page 118
∂R	The vertex boundary of region R, Page 8
$\pi_{\mathcal{B}}$ (vertices)	The finite-volume Gibbs distribution on Ω_R^+ under $\mathcal{B}$ of ∂R, Page 13
π_B (edges)	The finite-volume Gibbs distribution on Ω_R^+ induced by B of $\mathcal{E}R$, Page 51
$\pi_{\mathcal{B},\Lambda}$	The distribution on Ω_R^+ induced by $\pi_{\mathcal{B}}$, Page 13
π_k^σ	The uniform distribution on Ω_k^σ, Page 118
$p_X^{\min}(c,c')$	The probability v_X is coloured c and c' under $\Psi_X^{\min}$, Page 61
Ψ	A coupling of two distributions on colourings, Page 15
$\Psi_k(\sigma_1,\sigma_2)$	A coupling of $\pi_k^{\sigma_1}$ and $\pi_k^{\sigma_2}$, Page 118
$\Psi_X^{\min}$	A coupling of π_{B_X} and $\pi_{B'_X}$ minimising $\mathbb{E}[1_{\Psi,v_X}]$, Page 61
$p_v(\Psi_k(\sigma_1,\sigma_2))$	The probability $v \in \Theta_k$ is coloured differently under $\Psi_k(\sigma_1,\sigma_2)$, Page 118
Q	The set $\{1,\ldots,q\}$ of colours, Page 8
q	The number of colours, Page 7
Q_0	The set $\{0,1,\ldots,q\}$ of colours, Page 8
$\mathcal{R}$	An extended region, Page 80
$\mathbb{R}$	The set of real numbers, Page 7

R	A region (finite non-empty subset of $V_{\mathcal{G}}$), Page 8
R_v^d	The vertices $R \cap \mathrm{Ball}_d(v)$, Page 37
$\rho_{u,v}^k$	The influence of u on v under Θ_k, Page 118
R_X	The region of the edge-boundary pair X, Page 52
$R_{\mathcal{X}}$	The region of the vertex-boundary pair $\mathcal{X}$, Page 50
S	An $(\mathcal{A}, \mathcal{F})$-set, Page 77
σ, θ	Colourings of regions or graphs, Page 8
$\sigma(V)$	The colouring of vertices V induced by σ, Page 8
$\sigma(v)$	The colour of vertex v induced by σ, Page 8
$\tau_{\mathcal{M}}(\delta)$	The mixing time of $\mathcal{M}$, Page 16
Θ_k	They kth set of vertices to be updated in $\mathcal{M}_{\text{scan}}$, Page 19
T_X	The recursive-coupling tree for edge-boundary pair X, Page 61
$V_{\mathcal{G}}$	The vertex set of $\mathcal{G}$, Page 7
v_X	The endpoint in R_X of edge e_X in edge-boundary pair X, Page 52
w_X	The endpoint in $\mathcal{E}R_X$ of edge e_X in edge-boundary pair X, Page 52
$w_{\mathcal{X}}$	The distinguished boundary vertex in $\partial R_{\mathcal{X}}$, Page 50
$\mathcal{X}$	A vertex-boundary pair, Page 49
X	An edge-boundary pair, Page 52
$X_i(c, c')$	Edge-boundary pair constructed recursively in T_X, Page 60
$\mathbb{Z}$	The set of integers, Page 7
$Z_{\mathcal{B}}$	The partition function, Page 13
$\mathbb{Z}^+$	The set of positive integers, Page 7

Chapter 1

Introduction

Spin systems are a class of models that originated in Statistical physics. A *spin system* consists of a collection of *sites* which are the vertices of an underlying connected graph. The *spin space* is a finite set of *spins*, where a spin represents some physical property of a site. A *configuration* of the spin system is an assignment of a spin from the spin space to each site. The set of all configurations is referred to as the *configuration space*. The sites interact locally according to potentials specified by the spin system. This interaction gives rise to a well-defined probability distribution over configurations of any finite subset of the sites. Such a distribution is referred to as a *finite-volume Gibbs distribution* and is regarded as the equilibrium state of the given subset of the sites.

Probably the best known example of a spin system is the *Ising model*, introduced in the 1920's by Wilhem Lenz [Len20] and his student Ernst Ising [Isi25] as a simple model for magnetism. Let $G = (V, E)$ be a finite connected graph where each vertex $v \in V$ is a site. Each site can be seen as hosting a little magnet which is directed either up or down. Thus, the number of spins in the spin space is only two (spin up and spin down). The system has a certain temperature and we let $\beta \geq 0$ be the inverse temperature. There may be an external field characterised by a real value h. Given a configuration σ (an assignment of a spin to each vertex of G), let $m_=$ be the number of edges in G between two sites that are assigned the same spin, and let $m_{\neq}$ be the number of edges between sites that are assigned different spins. Let n_{up} be the number of sites with spin up, and let n_{down} be the number of sites with spin down. We now define the potential

$$H(\sigma) = \beta(m_{\neq} - m_=) + h\beta(n_{\text{up}} - n_{\text{down}}),$$

which is a function from the configuration space to the real numbers. The potential H depends on the interaction between neighbouring sites, as well as on sites individually when there is an external field ($h \neq 0$). The potential H is called the Hamiltonian and is defined differently for different spin systems. What characterises the Hamiltonian is

that it depends on local interactions between sites. When the system is in equilibrium, the probability of finding the system in configuration σ is proportional to $\exp(-H(\sigma))$. For the Ising model we therefore have that configurations in which many neighbouring sites are assigned the same spin are favoured by the Gibbs distribution. This effect increases with the inverse temperature β. That is, at high temperatures (low β) the spins behave almost independently, and at low temperatures (high β) there is a strong dependence between sites. Frequently, one imposes a *boundary condition* on the model, which corresponds to fixing the spin at some specified "boundary" vertices of G. The term *free boundary* is used to indicate that no boundary condition is specified.

In this thesis we focus on a much studied spin system known as the *anti-ferromagnetic Potts model*. It is similar to the Ising model above with the difference that we allow more than two spins, and adjacent sites that receive different spins are now favoured by the Gibbs distribution, whereas adjacent sites that receive the same spin are not favoured. As for the Ising model there are parameters representing the temperature and an external field. In this thesis we study the case with no external field. In the zero-temperature setting, a configuration in which two neighbouring sites have the same spin is not allowed. Such a configuration has zero probability under the Gibbs distribution. The tolerance towards equally-spinned neighbours is increasing with the temperature, and any configuration will have a positive probability at temperatures above zero.

Often the underlying graph of a spin system is infinite, in particular it is a lattice graph (an infinite graph for which the vertices are points that are regularly spaced). For a finite subset of the sites, referred to as a *region*, the boundary of a region R consists of the sites that are not in R but are incident to R. The finite-volume Gibbs distribution is defined for a region, whereas an *infinite-volume Gibbs distribution* is defined over configurations of the whole underlying graph. A distribution π on configurations of the underlying graph is an infinite-volume Gibbs distribution if, for any region R, the π-induced probability distribution on configurations of R, conditioned on a fixed assignment of spins to the boundary of R, is the finite-volume Gibbs distribution (under this fixed boundary). The physical intuition for an infinite-volume Gibbs distribution is that it describes a macroscopic equilibrium, for which all parts of the system are in equilibrium with their boundaries. There is always at least one infinite-volume Gibbs distribution, but in general, several infinite-volume Gibbs distributions can coexist [BW99]. The question of whether the infinite-volume Gibbs distribution is unique or not is central in statistical physics because it corresponds to the number of macroscopic equilibria for a given system. Non-uniqueness is known as a *phase transition*. For more background on Gibbs distributions, see for instance [Geo88] and [GHM01].

It is of interest to study what effect the boundary of a region has on sites that

are far away from it. For instance, if we change the spin of a site u on the boundary, then we ask what effect this change has (under the finite-volume Gibbs distribution) on the spin of a site v that is far away from u. If this effect is sufficiently small then it implies that the system has only one equilibrium (one unique infinite-volume Gibbs distribution). If the spins on the boundary has far-reaching effect on other sites, it might imply that there are several equilibria. We will study a concept called *strong spatial mixing*. A spin system that has strong spatial mixing has the property that the influence of a site u on the boundary dies out exponentially in the distance from u. If the system has strong spatial mixing then there is a unique equilibrium. In this thesis we analyse various spin systems under the anti-ferromagnetic Potts model in order to establish when there is strong spatial mixing.

We will present a general theorem that states sufficient conditions for when a system has strong spatial mixing under the anti-ferromagnetic Potts model. Whether there is strong spatial mixing or not depends on the maximum degree of the underlying graph as well as on the number of spins and the temperature. We let the parameter q denote the number of spins, and we let $\lambda \in [0, 1]$ denote the "temperature". More precisely, $\lambda = \exp(-\beta)$, where β is the inverse temperature, and we allow $\beta = \infty$ (zero-temperature) and $\beta = 0$ (infinitely high temperature).

Often the believed bounds for strong spatial mixing do not match the mixing bounds that have been rigorously proved. For instance, it is believed [SS97] that there is strong spatial mixing on the square lattice (Section 2.4.1) for every $q \geq 4$ and $\lambda \in [0, 1]$. However, for $q = 4$ and $q = 5$, strong spatial mixing has only successfully been shown for $\lambda \geq 0.262$ and $\lambda \geq 0.127$, respectively (this is proved in this thesis). The gap between the believed mixing bounds and the mixing bounds that follow from general theorems can often be narrowed by focusing on particular underlying graphs. Especially lattice graphs are suited for this because they enable us to make use of the regular geometry of the lattice. In this thesis we give proofs of improved mixing bounds for three types of lattices (see Figure 2.3 on page 24): the *square lattice* (Section 2.4.1), the *triangular lattice* (Section 2.4.2) and the *kagome lattice* (Section 2.4.3). These are all natural lattices that are of interest in statistical physics. Proofs of mixing that make use of the geometry of a particular graph sometimes depend on a great number of computations that are preferably (or necessarily) carried out with the help of a computer. Several of the proofs in this thesis are computer assisted.

Strong spatial mixing is a static property of a spin system. However, much attention has been on the dynamic properties of a system. A question of interest is to determine how quickly the system converges to equilibrium. From a statistical physics point of view, this question is important for understanding phenomena such as how long it takes for the system to return to equilibrium after a shock forces it out of it. In this thesis we consider a process called the *Glauber dynamics*, which is a Markov chain

on the set of configurations of a region. The *heat-bath* version of the Glauber dynamics works as follows. Suppose that R is a region and $\mathcal{B}$ is a fixed assignment of spins to the boundary of R. Let $\pi_{\mathcal{B}}$ be the finite-volume Gibbs distribution on configurations of R under the boundary $\mathcal{B}$. In one step of the Glauber dynamics we pick a vertex v of R uniformly at random, and replace the spin at v by a random spin drawn from the distribution on spins of v *conditional on* all the neighbouring spins. The stationary distribution of the Glauber dynamics is $\pi_{\mathcal{B}}$. The Glauber dynamics is a plausible model for the evolution of the underlying physical system towards equilibrium. Also, the Glauber dynamics can be used for sampling configurations from the Gibbs distribution. The concept of sampling is a major motivation for this thesis. Not only is sampling from the Gibbs distribution important in computational physics, but also there is a connection between sampling and *counting*. We will describe this connection shortly.

In order to use the Glauber dynamics as an efficient sampling tool we want to know how many steps it must be run for to get close to its stationary distribution. The notion of rate of convergence we study here is the *mixing time* of a Markov chain. The mixing time is defined as the number of steps required to get within total variation distance δ from the stationary distribution, starting from an arbitrary state of the Markov chain. If the mixing time of the Glauber dynamics is polynomial in the size of the region and $1/\delta$ then we say that it is *rapidly mixing*. It has been shown that strong spatial mixing and rapid mixing are two intimately related concepts. For example, [Ces01, DSVW04, Mar99, MO94a, MO94b, SZ92, Wei04] describe how strong spatial mixing implies rapid mixing in some settings. In this thesis we will describe in detail how the two concepts are related in our setting, and the results of strong spatial mixing are also presented as rapid-mixing results.

Recall from above that the probability of a configuration σ in the finite-volume Gibbs distribution π is proportional to $\exp(-H(\sigma))$, where H is the Hamiltonian for the spin system. That is,

$$\pi(\sigma) = \frac{\exp(-H(\sigma))}{Z},$$

where Z is a normalising factor that makes the sum of the probabilities over all configurations σ add up to one. The normalising factor Z is known as the *partition function*. In this thesis we will use a more graph theoretical terminology instead of the terminology that is commonly used by the physics community. This means that we will always refer to sites as vertices, and instead of using the word spin we will use the word *colour*. The spin space is therefore a set of colours, and a configuration is referred to as a *colouring*. A *proper colouring* is a colouring where no two adjacent vertices share the same colour. Thus, under the zero-temperature anti-ferromagnetic Potts model, the finite-volume Gibbs distribution is the uniform distribution on proper colourings. The partition function at zero temperature is therefore the number of proper colourings. Approximating

the partition function is a hard problem. In particular, the problem of counting proper colourings of a region is #P-complete [BDGJ99], where #P is the complexity class containing counting problems. For an introduction to #P, see for example [Jer03]. One implication if a #P-complete problem can be solved in polynomial time is that NP = P. Although exact counting of proper colourings is #P-complete, it does not rule out the possibility of *approximate counting*. If there is a known efficient method for uniformly sampling proper colourings then this method can be utilised in order to approximately count the number of proper colourings in an efficient way. For more background on the connection between sampling and counting, see for instance [Jer03, JVV86]. Therefore, if the Glauber dynamics mixes rapidly then it can be used (in a non-trivial way) to approximately count colourings in polynomial time.

At the end of the thesis we look at a dynamics referred to as a *systematic scan Markov chain*. It is similar to the Glauber dynamics and has gained increased attention recently ([DGJ06a, DGJ06b, Ped07]). As for the Glauber dynamics, the state space of the systematic scan Markov chain is the configuration space and the stationary distribution is the finite-volume Gibbs distribution. However, sites are updated in deterministic order instead of being chosen uniformly at random as in the Glauber dynamics. Here we analyse the systematic scan Markov chain on the square lattice under the zero-temperature anti-ferromagnetic Potts model, and we will present improved mixing bounds in this setting.

Organisation of the thesis

We summarise the outline of the thesis below.

Chapter 2 — Preliminaries. This chapter can be seen as a more detailed introduction to the thesis. We introduce notations and definitions and give examples of spin systems — in particular the anti-ferromagnetic Potts model. The concept of strong spatial mixing is formally defined and we explain in detail how sampling is achieved with the Glauber dynamics and the systematic scan Markov chain.

Chapter 3 — Statement of results. In this chapter we state the results that are proved in this thesis. Apart from Theorem 3.1 and 3.2, which have been published in [GJMP06], all other theorems represent new unpublished results that improve previously known mixing bounds. This chapter also covers some previous results on mixing which can be compared to the improvements presented in this thesis.

Chapter 4 — Rapidly mixing Glauber dynamics. In this chapter we explain the technique we use to show that the Glauber dynamics is rapidly mixing. We will see how the two concepts of rapid mixing and strong spatial mixing are related.

Chapter 5 — Boundary pairs. Here we introduce the framework we use when proving the mixing results in this thesis. This chapter introduces notation and definitions that were first used by Goldberg, Martin and Paterson in [GMP04], however we have modified them slightly to suit the contents of this thesis.

Chapter 6 — Recursive coupling. In order to prove the mixing theorems, we construct couplings of Gibbs distributions. A coupling is a joint distribution such that the marginal distributions are preserved. Using the framework from the previous chapter, this chapter explains how these couplings are constructed. We prove a general mixing result here and introduce some notation and lemmas that will be used in subsequent chapters when exploiting the geometry of the lattices in order to achieve better mixing bounds.

Chapter 7 — The square lattice. In this chapter we prove the mixing results for the square lattice.

Chapter 8 — The triangular lattice. In this chapter we prove the mixing results for the triangular lattice.

Chapter 9 — The kagome lattice. In this chapter we prove the mixing results for the kagome lattice.

Chapter 10 — Systematic scan on the grid. In this chapter we prove the rapid-mixing result with the systematic scan Markov chain on the square lattice.

Chapter 11 — Conclusion. Here we give thoughts on future work and discuss improvements of the results in this thesis.

Chapter 2

Preliminaries

In this chapter we give an introduction to the thesis that is more detailed and technical than in the previous chapter. We give definitions of all concepts needed in order to fully understand the theorems in Chapter 3 — Statement of results.

The organisation of this chapter is as follows. In Section 2.1 we introduce some basic notation and definitions, and we define the finite- and infinite-volume Gibbs distributions. We give a few examples of common spin systems, in particular the anti-ferromagnetic Potts model which is the spin system we focus on in this thesis. In Section 2.2 we give a formal definition of strong spatial mixing and define the notion of a K-coupling cover, which is related to strong spatial mixing. In Section 2.3 we define the Glauber dynamics and the systematic scan Markov chain, which are both used for sampling colourings. We give the definition of a rapidly mixing Markov chain and present a theorem which says that the Glauber dynamics is converging quickly to the desired distribution (i.e., it is rapidly mixing) if the system has a K-coupling cover. We also discuss the connection between efficiently sampling and approximating the partition function, or counting colourings. Lastly, in Section 2.4 we define what it means when an infinite graph is uniformly sub-exponential. The proofs of rapid mixing in this thesis require that the graph is neighbourhood-amenable, which is a property that follows if the graph is uniformly sub-exponential. In this section we define the square lattice, the triangular lattice and the kagome lattice, which are the lattices we focus on in Chapter 7–10. We show that they are uniformly sub-exponential.

2.1 Spin systems

A *spin system* is defined by a set of parameters: a connected, undirected, finite or infinite graph $\mathcal{G} = (V_\mathcal{G}, E_\mathcal{G})$, an integer $q \geq 1$ representing the number of *colours*, and a function H referred to as the *Hamiltonian*. The Hamiltonian describes the interaction between colours of vertices in $\mathcal{G}$. In the literature (especially the physics literature),

vertices of $\mathcal{G}$ are referred to as *sites* and colours are referred to as *spins*. In this thesis we will use the terms vertices and colours. We assume that the graph $\mathcal{G}$ is *locally finite*, which mean that the each vertex has a finite number of neighbours. We use Δ to denote the maximum degree of $\mathcal{G}$. Two vertices u and v are *adjacent* if (u, v) is an edge in $E_{\mathcal{G}}$. We let $d(u, v)$ be the distance between vertex u and v in $V_{\mathcal{G}}$, and we let $d(u, V)$ be the distance between vertex u and the subset $V \subseteq V_{\mathcal{G}}$. For an edge $e = (u_1, u_2) \in E_{\mathcal{G}}$, we let $d(e, v)$ be the minimum of $d(u_1, v)$ and $d(u_2, v)$.

A *region* R is a *finite non-empty* subset of $V_{\mathcal{G}}$. The subset $\partial R \subseteq V_{\mathcal{G}}$ denotes the *vertex boundary* of R such that ∂R is the set of vertices that are not in R but are adjacent to some vertex in R. The edge set $E(R)$ of R is the set of all edges $(u, v) \in E_{\mathcal{G}}$ such that at least one of the vertices u and v is in R. The *edge boundary* $\mathcal{E}R \subseteq E(R)$ of R is the set of all edges $(u, v) \in E_{\mathcal{G}}$ such that exactly one of the vertices u and v is in R and the other one is in ∂R.

The set $Q = \{1, \ldots, q\}$ denotes the set of q colours, and the set $Q_0 = \{0\} \cup Q$ denotes the set of q colours with the additional colour 0. The colour 0 represents "no colour". A *q-colouring* of a region R is a function from R to the set Q, and a q_0-colouring of R is a function from R to Q_0. A 0-colouring of R is a function from R to the set $\{0\}$, which means that all vertices in R are assigned colour 0. We often write only colouring when it is obvious from the context if it is a q-, q_0- or 0-colouring, or if any colouring will do. Let σ be a colouring of a region R. If R' is a subset of R then $\sigma(R')$ is the colouring of R' induced by σ. Furthermore, $\sigma(v)$ is the colour of vertex $v \in R$ under σ. The set Ω^+ denotes all q-colourings of $V_{\mathcal{G}}$ and Ω_R^+ denotes all q-colourings of the region R. A colouring is *proper* if no two adjacent vertices receive the same colour. The set Ω denotes all proper q-colourings of $V_{\mathcal{G}}$ and Ω_R denotes all proper q-colourings of the region R. Given a q_0-colouring $\mathcal{B}$ of ∂R, a proper q-colouring σ of R *agrees* with $\mathcal{B}$ if $\sigma(u) \neq \mathcal{B}(v)$ for all $(u, v) \in \mathcal{E}R$, where $u \in R$. The set $\Omega_R(\mathcal{B})$ denotes the set of all proper q-colourings of R that agree with $\mathcal{B}$. For two colourings $\sigma, \sigma' \in \Omega_R^+$, the *Hamming distance* between σ and σ' is the number of vertices in R on which σ and σ' differ. In the physics literature the set Q of colours is often referred to as the *state space* or *spin space*, and a colouring is referred to as a *configuration*. The set Ω^+ of colourings is referred to as the *configuration space*.

When illustrating graphs or regions in this thesis we denote vertices with faces instead of drawing the graph in the traditional way where a vertex is denoted with a solid circle and an edge is denoted with a line segment. Using faces gives us space to label vertices or write other useful information. See Figure 2.1 for an example. When a vertex is labelled with a number, this number is referring to the colour of the vertex.

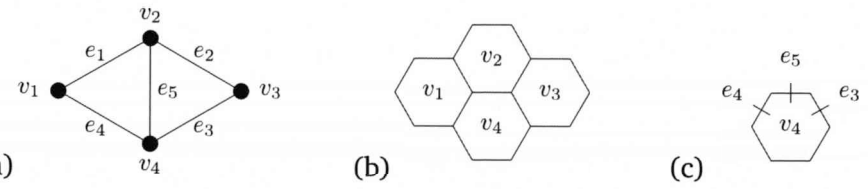


Figure 2.1. We illustrate a graph (a) such that faces correspond to vertices (b). Short line segments represent edges (c).

2.1.1 Gibbs distributions

In this thesis we consider spin systems with nearest-neighbour interactions only. A colouring is assigned a weight, and the weight depends on the colours assigned to adjacent vertices. We define the Hamiltonian as follows. The definition is general and is based on parameters that are specific for different spin systems.

Definition 2.1 (Hamiltonian). *For any region R and any q-colouring σ of the whole graph $\mathcal{G}$, the* Hamiltonian

$$H_R(\sigma) = \sum_{(u,v)\in E(R)} U(\sigma(u),\sigma(v)) + \sum_{v\in R} W(\sigma(v)),$$

where $U : Q \times Q \to \mathbb{R} \cup \{\infty\}$ is a symmetric function that maps a pair of colours to a real value or infinity, and $W : Q \to \mathbb{R} \cup \{\infty\}$ is a function that maps a colour to a real value or infinity.

The Hamiltonian assigns a value to each colouring σ of a region R. This value can in physics be interpreted as the contribution to the energy from the region R given by the configuration σ. What defines the spin system is specified with the functions U and W. The function U describes the weight contributed by neighbouring vertices, and W describes the weight contributed by a single vertex. In physics, U is the potential associated with neighbouring sites, and W is associated with a self potential which might come from the action of an external field. The Hamiltonian is used in the definition of a *finite-volume Gibbs distribution* below. For details and more on Gibbs distributions, see [Geo88] or [GHM01].

Definition 2.2 (Finite-volume Gibbs distribution). *For any region R and any q-colouring η of the whole graph $\mathcal{G}$, the* finite-volume Gibbs distribution π_R^η *on the set Ω^+ of all q-colourings of $\mathcal{G}$ is*

$$\pi_R^\eta(\sigma) = \begin{cases} \frac{1}{Z_R^\eta}\exp(-H_R(\sigma)) & \text{if } \sigma(\mathcal{G}\backslash R) = \eta(\mathcal{G}\backslash R); \\ 0 & \text{otherwise,} \end{cases}$$

where Z_R^η is a normalising factor that makes $\sum_{\sigma\in\Omega^+}\pi_R^\eta(\sigma) = 1$.

In the finite-volume Gibbs distribution π_R^η, a colouring σ of $\mathcal{G}$ cannot have a positive probability unless σ is identical to η on all vertices outside of R. The probability of choosing a colouring σ from π_R^η that is identical to η outside of R is proportional to $\exp(-H_R(\sigma))$. Since the Hamiltonian depends only on neighbour interactions and single vertices in R, it is sufficient to know the colouring of ∂R under η in order to specify the finite-volume Gibbs distribution. The normalising factor Z_R^η is referred to as the *partition function*. Approximating the partition function is sometimes desirable and will be discussed in Section 2.3.3.

After defining finite-volume Gibbs distributions we define *infinite-volume Gibbs distributions* in terms of conditional probabilities.

Definition 2.3 (Infinite-volume Gibbs distribution). *A distribution π on the set Ω^+ of all q-colourings of $\mathcal{G}$ is called an* infinite-volume Gibbs distribution *if for all regions R, all colourings $\sigma \in \Omega^+$ and $\eta \in \Omega^+$ with $\pi(\eta) > 0$,*

$$\Pr[C(R) = \sigma(R) \,|\, C(\mathcal{G} \backslash R) = \eta(\mathcal{G} \backslash R)] = \pi_R^\eta(\sigma),$$

where C is a colouring drawn from π.

Thus, if a colouring C is drawn from π with the vertices outside of region R coloured identical to η, then the conditional probability that the colouring of the vertices in R is identical to σ is $\pi_R^\eta(\sigma)$. This definition of an infinite-volume Gibbs probability goes back to the work of Dobrushin [Dob68] and Lanford and Ruelle [LR69] in the late 1960's. The physical intuition for an infinite-volume Gibbs distribution is that it describes a macroscopic equilibrium, for which all parts of the system are in equilibrium with their boundaries. There is always at least one infinite-volume Gibbs distribution, but in general, several infinite-volume Gibbs distributions for the same Hamiltonian can coexist [BW99]. The question of whether the infinite-volume Gibbs distribution is unique or not is central in statistical physics because it corresponds to the number of macroscopic equilibria for a given system. The phenomenon of non-uniqueness corresponds to what is referred to as a *phase transition*. For more on Gibbs distributions and details on the connection with physics, see [Geo88] or [GHM01]. For more on different criteria for uniqueness, see for example [DSVW04, Wei04, Wei05]. The following example illustrates a spin system for which an infinite number of infinite-volume Gibbs distributions exist.

Example 2.4 (Many infinite-volume Gibbs distributions). Consider the spin system defined by $q = 2$ colours on the square lattice $\mathcal{G} = \mathbb{Z}^2$. The is Hamiltonian defined by

$$U(c, c') = \begin{cases} \infty & \text{if } c = c'; \\ 0 & \text{if } c \neq c', \end{cases} \qquad \text{and} \qquad W(c) = 0,$$

where $c, c' \in Q$. Let η_1 and η_2 be the two different colourings of $\mathcal{G}$ that colour the vertices like a chessboard. That is, each vertex has a colour that is different from its four neighbours. The colourings η_1 and η_2 are each other's inverse. Let π be the distribution on Ω^+ that assigns probability p to η_1 and probability $1-p$ to η_2, where $p \in [0,1]$ is chosen arbitrarily. All other colourings in Ω^+ are assigned zero probability. It can now be verified that π satisfies the condition of being an infinite-volume Gibbs distribution. To see this, consider any region R. If $\eta = \eta_1$ then for $\sigma = \eta_1$ we have that both the conditional probability in Definition 2.3 and $\pi_R^\eta(\sigma)$ are 1. For $\sigma \neq \eta_1$ we have that the conditional probability and $\pi_R^\eta(\sigma)$ are both zero. Similarly, if $\eta = \eta_2$ then for $\sigma = \eta_2$ we have that both probabilities are 1. For $\sigma \neq \eta_2$ we have that both probabilities are zero. Thus, π is an infinite-volume Gibbs distribution. Since the probability p can be chosen arbitrarily there is an infinite number of infinite-volume Gibbs distributions π.

Next we will give a few examples of some very common spin systems that have been studied extensively in the literature. one of them is the *anti-ferromagnetic Potts model*, which is the spin system we focus on in this thesis.

Example 2.5 (The ferromagnetic Ising model)**.** Probably the best known spin system is the *Ising model*, introduced in the 1920's by Wilhem Lenz [Len20] and his student Ernst Ising [Isi25] as a simple model for magnetism. The number of colours $q = 2$, representing two spins: up and down. The Hamiltonian is given by

$$U(c,c') = \begin{cases} 0 & \text{if } c = c'; \\ \beta & \text{if } c \neq c' \end{cases} \qquad \text{and} \qquad W(c) = \begin{cases} h\beta & \text{if } c = 1; \\ -h\beta & \text{if } c = 2, \end{cases}$$

where $c, c' \in Q$ and $\beta \geq 0$ is the inverse temperature and $h \in \mathbb{R}$ is an external field. The infinite-volume Gibbs distribution assigns higher weight to colourings in which many adjacent vertices receive the same colour, as well as to colourings in which many vertices receive the colour that is favoured by the sign of h. At high temperatures (low β), the colours of adjacent vertices have a smaller impact on the weight of a colouring. At low temperatures (high β), the colours of adjacent vertices become more important and large connected regions of equal colours tend to form. The higher the magnitude $|h|$ of the external field, the less symmetric is the model in the two colours. If $|h| = 0$ there is an equal chance of having many adjacent vertices with colour 1 or having many adjacent vertices with colour 2.

Example 2.6 (The anti-ferromagnetic Ising model)**.** The *anti-ferromagnetic Ising model* is very similar to the Ising model, only with the difference that the sign before β in the function U has changed. This means that adjacent vertices now prefer to take opposite colours.

Example 2.7 (The hard-core model). The *hard-core model* (independent sets) has been used in statistical physics as a model of lattice gases [Geo88]. The number of colours $q = 2$, where a vertex with colour 1 is *occupied* and a vertex with colour 2 is *unoccupied*. A colouring therefore specifies a subset of occupied vertices. The Hamiltonian is given by

$$U(c, c') = \begin{cases} \infty & \text{if } c = c' = 1; \\ 0 & \text{otherwise} \end{cases} \qquad \text{and} \qquad W(c) = \begin{cases} -\ln \ell & \text{if } c = 1; \\ 0 & \text{if } c = 2, \end{cases}$$

where $\ell > 0$ is referred to as the activity parameter and plays a similar role to that of temperature in the Ising model. The Hamiltonian forbids two neighbouring vertices from both being occupied. Thus, in this model, a colouring is valid if, and only if, it specifies an *independent set*. The finite-volume Gibbs distributions are over independent sets and the probability of an independent set σ is proportional to $\ell^{|\sigma|}$. For low values of ℓ (that is $\ell < 1$) the density of occupied vertices is low. On the other hand, for large values of ℓ (that is $\ell > 1$) the density of occupied vertices is high, which means that the constraint of not having two adjacent occupied vertices has a greater impact on the distributions of colourings. Weitz shows in [Wei06] uniqueness of the infinite-volume Gibbs distribution for $\ell < (\Delta - 1)^{\Delta-1}/(\Delta - 2)^{\Delta}$, where Δ is the maximum degree of $\mathcal{G}$. This is an improvement of the general result $\ell \leq 2/(\Delta - 2)$ by Dyer and Greenhill [DG00] and Vigoda [Vig01].

Example 2.8 (The ferromagnetic Potts model). A natural generalisation of the ferromagnetic Ising model is the *ferromagnetic Potts model* [Pot52], in which the number of colours q may be more than two. Hence the Ising model is the special case when $q = 2$. For the case without an external field we have that the Hamiltonian is given by

$$U(c, c') = \begin{cases} 0 & \text{if } c = c'; \\ \beta & \text{if } c \neq c' \end{cases} \qquad \text{and} \qquad W(c) = 0,$$

where $c, c' \in Q$ and $\beta \geq 0$ is the inverse temperature. As for the Ising model the system favours colourings with many adjacent vertices receiving the same colour. In this thesis we focus on the anti-ferromagnetic Potts model, which is described in the next section.

2.1.2 The anti-ferromagnetic Potts model

The *anti-ferromagnetic Potts model* is similar to the ferromagnetic Potts model, only with the difference that different colours on adjacent vertices are favoured. The spin system is specified by $q \geq 2$ number of colours and an inverse temperature $\beta \geq 0$. The

Hamiltonian is given by

$$U(c,c') = \begin{cases} \beta & \text{if } c = c'; \\ 0 & \text{if } c \neq c' \end{cases} \qquad \text{and} \qquad W(c) = 0,$$

where $c, c' \in Q$. Instead of explicitly using the inverse temperature β we will use the parameter $\lambda = \exp(-\beta)$, where $\lambda \in [0,1]$ when allowing $\beta = \infty$ (zero-temperature) and $\beta = 0$ (infinitely high temperature). The zero-temperature case ($\lambda = 0$) is particularly interesting since it corresponds to proper colourings of the graph.

Consider any region R and let σ be any colouring of R and let $\mathcal{B}$ be any colouring of ∂R. An edge $(u,v) \in E(R)$ is *monochromatic* if u and v receive the same colour under σ and $\mathcal{B}$. For a set $E \subseteq E(R)$ of edges, let $\text{mon}_{\sigma,\mathcal{B}}(E)$ denote the number of monochromatic edges in E under σ and $\mathcal{B}$. A proper colouring is a colouring with no monochromatic edges. Using the notion of monochromatic edges and the parameter λ above, the finite-volume Gibbs distribution from Definition 2.2 is given by

$$\pi_R^\eta(\sigma) = \begin{cases} \frac{1}{Z_R^\eta} \lambda^{\text{mon}_{\sigma(R),\eta(\partial R)}(E(R))} & \text{if } \sigma(\mathcal{G}\backslash R) = \eta(\mathcal{G}\backslash R); \\ 0 & \text{otherwise,} \end{cases}$$

where Z_R^η is a normalising factor such that

$$\sum_{\sigma \in \Omega^+} \pi_R^\eta(\sigma) = 1.$$

The distribution on Ω_R^+, all q-colourings of R, induced by the finite-volume Gibbs distribution, will be denoted $\pi_{\mathcal{B}}$ and is defined by

$$\pi_{\mathcal{B}}(\sigma) = \frac{\lambda^{\text{mon}_{\sigma,\mathcal{B}}(E(R))}}{Z_{\mathcal{B}}},$$

where $\mathcal{B}$ is a colouring of the boundary ∂R, and

$$Z_{\mathcal{B}} = \sum_{\sigma \in \Omega_R^+} \lambda^{\text{mon}_{\sigma,\mathcal{B}}(E(R))}$$

is the partition function. For any subset $\Lambda \subseteq R$, let $\pi_{\mathcal{B},\Lambda}$ denote the distribution on colourings in Ω_Λ^+ induced by $\pi_{\mathcal{B}}$.

A proper colouring σ of R *agrees* with a boundary colouring $\mathcal{B}$ of ∂R if no edges in $E(R)$ are monochromatic under σ and $\mathcal{B}$. Note that for $\lambda = 0$, the distribution $\pi_{\mathcal{B}}$ is the uniform distribution on proper colourings of R that agree with the colouring $\mathcal{B}$ of the boundary.

Remark 2.9. In this thesis, we will always assume that $q \geq \Delta + 1$ in the zero-temperature

setting. The reason for this is that it will guarantee that for any region R and any q-colouring $\mathcal{B}$ of ∂R, there is at least one proper colouring σ of R that agrees with $\mathcal{B}$. Thus, $\pi_{\mathcal{B}}$ is well-defined with non-empty support. Note that a greedy colouring of R would yield a proper colouring that agrees with $\mathcal{B}$. That is, for each vertex of R, assign a colour that is not assigned to any of its neighbours. This is always possible since the number of colours $q \geq \Delta + 1$.

2.2 Strong spatial mixing

The concept of strong spatial mixing is central in this thesis and it has two important consequences. One consequence is that *strong spatial mixing* (Definition 2.10) often implies that there is a unique infinite-volume Gibbs distribution (see [Mar99, Wei04]). The second consequence is that strong spatial mixing implies that the *Glauber dynamics* (Definition 2.14) is *rapidly mixing*. The Glauber dynamics, introduced in Section 2.3.1, is a Markov chain used for sampling colourings of a region R.

Informally, strong spatial mixing means that for any region R and two colourings $\mathcal{B}$ and $\mathcal{B}'$ of ∂R that differ only at one vertex $w \in \partial R$, the effect that this difference has on a subset $\Lambda \subseteq R$ decays exponentially with the distance from w to Λ. The effect is measured with the *total variation distance*. For two distributions D_1 and D_2 on a set S, the total variation distance between D_1 and D_2 is defined as

$$\mathrm{d_{TV}}(D_1, D_2) = \frac{1}{2}\sum_{s\in S}|D_1(s) - D_2(s)| = \max_{A\subseteq S}|D_1(A) - D_2(A)|.$$

The following formal definition of strong spatial mixing is taken from [DSVW04].

Definition 2.10 (Strong spatial mixing)**.** *The anti-ferromagnetic Potts model on the infinite graph $\mathcal{G}$ has* strong spatial mixing *for parameters q and λ if there are two constants $\alpha > 0$ and $0 < \varepsilon < 1$ such that, for any region R, any $\Lambda \subseteq R$, any vertex $w \in \partial R$, and any pair $(\mathcal{B}, \mathcal{B}')$ of q_0-colourings of ∂R which differ only at w,*

$$\mathrm{d_{TV}}(\pi_{\mathcal{B},\Lambda}, \pi_{\mathcal{B}',\Lambda}) \leq \alpha|\Lambda|(1-\varepsilon)^{d(w,\Lambda)},$$

where $d(w, \Lambda)$ is the distance within R from the vertex w to the region Λ. We assume that w is not a free-boundary vertex in either colouring. That is, $\mathcal{B}(w) \neq 0$ and $\mathcal{B}'(w) \neq 0$.

Note that strong spatial mixing is defined for q_0-colourings of the boundary of the region. The reason we allow the colour 0 (interpreted as "no colour") on the boundary is to strengthen the strong spatial mixing results. This will be important further on when we derive techniques for sampling colourings. For a region R and parameters q and λ, we want to sample colourings from $\pi_{\mathcal{B}}$, where $\mathcal{B}$ is a q_0-colouring of ∂R. In

particular, under no boundary condition, $\mathcal{B}$ is the 0-colouring (free boundary). To analyse sampling under the free boundary, it will be important to include the colour 0 in the strong spatial mixing results. Chapter 4 explains the relationship between strong spatial mixing and sampling in detail.

One approach to show exponential decay of the total variation distance in the distance between w and Λ in Definition 2.10 is to construct a suitable *coupling* of the distributions $\pi_{\mathcal{B}}$ and $\pi_{\mathcal{B}'}$. For two distributions D_1 and D_2 on a set S, a coupling Ψ of D_1 and D_2 is a joint distribution on $S \times S$ with marginal distributions D_1 and D_2. If the pair (X_1, X_2) is a random variable drawn from Ψ then

$$\mathrm{d_{TV}}(D_1, D_2) \leq \Pr(X_1 \neq X_2).$$

Thus, in order to upper-bound the total variation distance, one can find some suitable coupling Ψ and compute the the probability of having $X_1 \neq X_2$. In the case of strong spatial mixing in Definition 2.10, the aim is to construct a coupling Ψ of $\pi_{\mathcal{B}}$ and $\pi_{\mathcal{B}'}$ such that if the pair (σ, σ') of colourings is drawn from Ψ, then the probability that σ and σ' differ on $\Lambda \subseteq R$ decreases exponentially with the distance between the discrepancy vertex $w \in \partial R$ and Λ. If such a coupling Ψ can be constructed then strong spatial mixing follows (Lemma 5.4). Furthermore, if the graph $\mathcal{G}$ is of sub-exponential growth (see Section 2.4) then it follows that the expected number of vertices in R on which σ and σ' differ is bounded by a constant (Lemma 5.5). We will make use of this fact when showing rapid mixing of the Glauber dynamics. Next we introduce the notion of a *K-coupling cover*, which is equivalent to Goldberg, Martin and Paterson's notion of an ε-coupling cover in [GMP05, Definition 4].

Definition 2.11 (K-coupling cover). *The anti-ferromagnetic Potts model on an infinite graph $\mathcal{G}$ with parameters q and λ has a* K-coupling cover *if, for every region R and any pair $(\mathcal{B}, \mathcal{B}')$ of q_0-colourings of ∂R, which differ only at a vertex $w \in \partial R$ (and $\mathcal{B}(w) \neq 0$ and $\mathcal{B}'(w) \neq 0$), there is a coupling Ψ of $\pi_{\mathcal{B}}$ and $\pi_{\mathcal{B}'}$ with the following property: If the pair (σ, σ') of colourings is drawn from Ψ then the expected number of vertices in R on which σ and σ' differ is bounded by the constant K.*

Note that the constant K depends on the graph $\mathcal{G}$, the number of colours q and the parameter λ.

2.3 Sampling

In previous sections we described what an equilibrium state is in terms of the finite- and infinite-volume Gibbs distributions. A question of interest is to determine how quickly the system converges to equilibrium. From a statistical physics point of view,

this question is important for understanding phenomena such as how long it takes for the system to return to equilibrium after a shock forces it out of it. In this thesis we consider two dynamical processes that model how the system converges. Both processes are *Markov chains* which are used to sample from finite-volume Gibbs distributions. Under the anti-ferromagnetic Potts model, for a given region R and colouring $\mathcal{B}$ of the boundary ∂R, these processes can be used for sampling colourings from the distribution $\pi_{\mathcal{B}}$. The notion of rate of convergence we study in this thesis is the *mixing time* of a Markov chain, which is a common notion of convergence (see for instance [Jer03]). The mixing time is defined as the number of steps required to get within total variation distance δ from the stationary distribution, starting from an arbitrary state of the Markov chain. The formal definition is given next.

Definition 2.12 (Mixing time). *Consider a Markov chain $\mathcal{M}$ with state space Ω and stationary distribution π. Let $P^t(\sigma, \sigma')$ be the probability of moving from state σ to σ' in exactly t steps. For any $\delta > 0$, the* mixing time *of $\mathcal{M}$ is*

$$\tau_{\mathcal{M}}(\delta) = \max_{\sigma \in \Omega} \min\{t : \mathrm{d_{TV}}(P^t(\sigma, \cdot), \pi_{\mathcal{B}}) \leq \delta\}.$$

A Markov chain $\mathcal{M}$ that is used for sampling colourings from the distribution $\pi_{\mathcal{B}}$, given a region R and a colouring $\mathcal{B}$ of ∂R, is said to be *rapidly mixing* if $\tau_{\mathcal{M}}(\delta)$ is upper-bounded by a polynomial in n and $\log(1/\delta)$, where n is the number of vertices in R.

The main dynamical process we focus on in this thesis is called the (*heat-bath*) *Glauber dynamics*. A step in this Markov chain is a random update of the colour of a single vertex, depending on the colours of its neighbours. The vertex to be updated is chosen uniformly at random from R. As we will see, this process converges to the distribution $\pi_{\mathcal{B}}$.

The other process we look at, referred to as a (*heat-bath*) *systematic scan Markov chain*, is similar to the Glauber dynamics, but instead of choosing the vertex to update uniformly at random, all vertices in R are randomly updated in deterministic order. Rather than updating a single vertex, often a subset of the vertices of R is updated, conditional on the colours of the boundary of the subset. A dynamics that updates more than one vertex is referred to as a *block dynamics*. As we will see, the systematic scan Markov chain also converges to the distribution $\pi_{\mathcal{B}}$.

In the literature, the term "Glauber dynamics" usually refers to any Markov chain that makes single-vertex updates. Here we use the term Glauber dynamics exclusively for the first Markov chain described above. That is, the process in which a vertex is chosen uniformly at random.

The term "heat-bath" means that the update of a vertex v (or set V of vertices) is done according to the stationary distribution on the colours of v (or V), conditional

on the colours of the neighbours of v (or the boundary of V). A *heat-bath update* is defined as follows.

Definition 2.13 (Heat-bath update). *For any region R, any q_0-colouring $\mathcal{B}$ of ∂R, any q_0-colouring σ of R and any non-empty subset $\Theta \subseteq R$, a* heat-bath update of σ with respect to Θ *is a procedure in which a new colouring σ' of R is obtained from σ by recolouring the vertices in Θ in the following way: Let $\mathcal{B}_{\sigma,\Theta}$ be the colouring of $\partial\Theta$ induced by σ and $\mathcal{B}$. Obtain the new colouring σ' by drawing a colouring from $\pi_{\mathcal{B}_{\sigma,\Theta}}$ and assign that colouring to the vertices in Θ. Each vertex in $R\backslash\Theta$ is assigned the colour it has in σ.*

2.3.1 The Glauber dynamics

We have the following definition of the Glauber dynamics.

Definition 2.14 (Glauber dynamics). *For any region R and any q_0-colouring $\mathcal{B}$ of ∂R, the* Glauber dynamics *is a Markov chain defined as follows.*

- *The state space is Ω_R^+ if $\lambda > 0$ and $\Omega_R(\mathcal{B})$ if $\lambda = 0$.*
- *A transition is made from a state σ to σ' in the following way:*
  1. *Choose a vertex v uniformly at random from R.*
  2. *For $i \in Q$, let n_i denote the number of neighbours of v which are assigned colour i (either in σ or $\mathcal{B}$).*
  3. *Choose a new colour $c \in Q$ according to the distribution*

$$\Pr[c = i] = \frac{\lambda^{n_i}}{\sum_{k \in Q} \lambda^{n_k}}$$

  for $i \in Q$ and with $0^0 = 1$.
  4. *Obtain the new colouring σ' from σ by assigning colour c to vertex v.*

The reason for defining the Glauber dynamics on two different state spaces (Ω_R^+ if $\lambda > 0$ and $\Omega_R(\mathcal{B})$ if $\lambda = 0$) is to avoid colourings that are not proper or do not agree with the boundary colouring $\mathcal{B}$ when $\lambda = 0$. Such colourings have zero probability in $\pi_{\mathcal{B}}$ under the anti-ferromagnetic Potts model. If the Glauber dynamics is ergodic (aperiodic and irreducible) then the distribution $\pi_{\mathcal{B}}$ is the stationary distribution (for a proof, see Lemma 4.2).

The Glauber dynamics is ergodic for positive values of λ; each state can be reached from any state and there is always a positive probability of not leaving a state. For $\lambda = 0$ this is not always the case. However, if the number of colours $q \geq \Delta + 2$ then Glauber dynamics is ergodic. To see this, take any two colourings σ and σ' in $\Omega_R(\mathcal{B})$

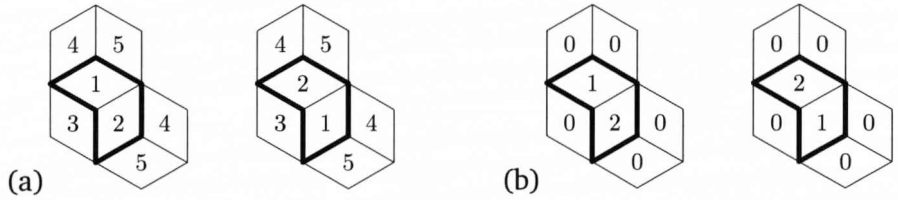


Figure 2.2. A 2-vertex region R of the kagome lattice. The Glauber dynamics on R is not irreducible for $\lambda = 0$ and $q = 5$ under the colouring of ∂R shown in (a). The two states in (a) are "frozen". However, when restricting the boundary to the 0-colouring as in (b), the Glauber dynamics is irreducible.

and let $v_1, \ldots, v_m$ be the vertices on which σ and σ' differ. To go from state σ to σ' we update the vertices $v_1, \ldots, v_m$ in order. The vertex v_i is updated as follows: If any neighbours of v_i have colour $\sigma'(v_i)$, recolour these with the first available colour. They can always be recoloured since $q \geq \Delta + 2$. Note that these neighbours do not have their final colour in σ'. Once no neighbour of v_i has colour $\sigma'(v_i)$, recolour v_i with colour $\sigma'(v_i)$.

The Glauber dynamics is not necessarily ergodic on the set of proper colourings if $q < \Delta + 2$ and $\lambda = 0$. Figure 2.2(a) illustrates two proper colourings of a 2-vertex region R of the kagome lattice (see Section 2.4.3) with a particular colouring $\mathcal{B}$ of ∂R. The Glauber dynamics is not irreducible on R under $\mathcal{B}$ when $q = 5$ and $\lambda = 0$. The two states in Figure 2.2(a) are "frozen" and it is not possible to leave either of them by single-vertex updates. However, as Figure 2.2(b) shows, if we restrict the colouring $\mathcal{B}$ to the 0-colouring then it is possible to go between any two states. The conclusion is that if $\lambda = 0$ and $q < \Delta + 2$ we need to pay extra attention when studying the Glauber dynamics.

Theorem 2.15 below, which is proved in Section 4, guarantees rapidly mixing Glauber dynamics if the spin system has a K-coupling cover and the parameter $\lambda > 0$. If $\lambda = 0$, rapid mixing is only guaranteed with the additional condition that $q \geq \Delta + 2$ (which implies that the Markov chain is always irreducible). In this thesis we also show that the Glauber dynamics is rapidly mixing on the kagome lattice ($\Delta = 4$) for $q = 5$ and $\lambda = 0$. However, here we restrict the boundary colouring of the region to be the free colouring (0-colouring) in order to make sure that the Markov chain is irreducible (see Theorem 3.5). The notion of a uniformly sub-exponential graph that is used in Theorem 2.15 is defined in Section 2.4. Note that the mixing bound $O(n(n + \log \frac{1}{\delta}))$ in the theorem is not optimal. This will be discussed at the end of Section 4.4.

Theorem 2.15. *Consider the anti-ferromagnetic Potts model on a uniformly sub-exponential graph $\mathcal{G}$ with maximum degree Δ, and parameter q and λ. If the spin system has a K-coupling cover, and $\lambda > 0$ or $q \geq \Delta + 2$, then the Glauber dynamics is rapidly mixing for any region R and any q_0-colouring $\mathcal{B}$ of ∂R. More precisely, the mixing time*

$\tau(\delta) \leq O(n(n + \log \frac{1}{\delta}))$, *where* n *is the number of vertices in* R.

2.3.2 Systematic scan

Previously we have defined the Glauber dynamics. A related dynamics that we study briefly in this thesis is the systematic scan Markov chain. The difference to the Glauber dynamics is that the systematic scan Markov chain updates vertices in deterministic order, as well as updates not necessarily a single vertex but several vertices in each step. The systematic scan Markov chain $\mathcal{M}_{\text{scan}}$ is defined as follows.

Definition 2.16 (Systematic scan Markov chain $\mathcal{M}_{\text{scan}}$). *For any region* R, *any* q_0*-colouring* $\mathcal{B}$ *of* ∂R *and any collection* $\Theta_1, \ldots, \Theta_m$ *of* m *subsets* $\Theta_k \subseteq R$ *such that* $\cup_{k=1}^{m} \Theta_k = R$, $m \in O(|R|)$ *and* $1 \leq |\Theta_k| \leq K$, *where* $k = 1, \ldots, m$ *and* K *is a constant, the* heat-bath systematic scan Markov chain $\mathcal{M}_{\text{scan}}$ *is defined as follows.*

- *The state space is* Ω_R^+ *if* $\lambda > 0$ *and* $\Omega_R(\mathcal{B})$ *if* $\lambda = 0$.
- *A transition is made from a state* σ *to* σ' *in the following way:*
  1. *Let* $\sigma_0 = \sigma$.
  2. *For* $k = 1, \ldots, m$ *(starting with* $k = 1$ *and then increasing* k *by 1 until* $k = m$*), obtain a new colouring* σ_k *by performing a heat-bath update of* σ_{k-1} *with respect to* Θ_k.
  3. *The new colouring* $\sigma' = \sigma_m$.

One could say that in general it has shown to be more difficult to analyse the mixing time of dynamics that update vertices in deterministic order compared to random updates as in the Glauber dynamics. That is, several spin systems have been shown to converge quickly to the Gibbs distribution when using a random update approach, whereas fast convergence using a systematic scan approach still remains to be proved. For results regarding systematic scan, see for instance Dyer, Goldberg and Jerrum [DGJ06a, DGJ06b] and Pedersen [Ped07].

As for the Glauber dynamics (see previous section) there could be problems with ergodicity of $\mathcal{M}_{\text{scan}}$ if the number of colours is too small. In this thesis we only consider settings where the number of colours is large enough to guarantee that $\mathcal{M}_{\text{scan}}$ is ergodic. As for the Glauber dynamics, the stationary distribution of $\mathcal{M}_{\text{scan}}$ is $\pi_{\mathcal{B}}$. For completeness we provide a proof for this fact. Since we only consider systematic scan Markov chains in the zero-temperature setting, the following lemma is stated with the restriction that $\lambda = 0$.

Lemma 2.17. *If* $\lambda = 0$ *and* $\mathcal{M}_{\text{scan}}$ *is ergodic, then the stationary distribution of* $\mathcal{M}_{\text{scan}}$ *is* $\pi_{\mathcal{B}}$.

Proof. For $k = 1, \ldots, m$, let P_k be the $|\Omega_R(\mathcal{B})| \times |\Omega_R(\mathcal{B})|$-transition matrix containing all transition probabilities under a heat-bath update with respect to Θ_k. Let $P_k(i, j)$ denote the probability of making a transition from colouring i to colouring j. Note that we refer to a colouring with the row index i and column index j of the transition matrix. We have that $P = \prod_{k=1}^{m} P_k$ is the transition matrix for the systematic scan Markov chain $\mathcal{M}_{\text{scan}}$. We will show that $\pi_{\mathcal{B}} P_k = \pi_{\mathcal{B}}$ for all $k = 1, \ldots, m$. Then it follows that $\pi_{\mathcal{B}} P = \pi_{\mathcal{B}}$, which means that $\pi_{\mathcal{B}}$ is a stationary distribution.

Fix a value of $k \in \{1, \ldots, m\}$. In order to show that $\pi_{\mathcal{B}} P_k = \pi_{\mathcal{B}}$ we will show that the row vector $\pi_{\mathcal{B}}$ multiplied by the jth column of P_k equals the jth element in $\pi_{\mathcal{B}}$, for all columns j. Thus, we will show that

$$\sum_i \pi_{\mathcal{B}}(i) P_k(i, j) = \pi_{\mathcal{B}}(j). \tag{2.1}$$

For a colouring j of R, let $\Omega_j \subseteq \Omega_R(\mathcal{B})$ be the set of colourings σ of R such that $\sigma(R \setminus \theta_k) = j(R \setminus \theta_k)$ and $\sigma(\Theta_k)$ is a proper colouring of Θ_k that agrees with the colouring of $\partial\Theta_k$ induced by j and $\mathcal{B}$. Then $P_k(i, j) = 0$ if $i \notin \Omega_j$ since the heat-bath update only recolours the vertices in Θ_k. If $i \in \Omega_j$ then, since $\lambda = 0$, the heat-bath update chooses any of the $|\Omega_j|$ colourings in Ω_j uniformly at random, and $P_k(i, j) = 1/|\Omega_j|$. Assume that there are N proper q-colourings of R that agree with $\mathcal{B}$. Then $\pi_{\mathcal{B}}(i) = 1/N$ for all $i \in \Omega_R(\mathcal{B})$. Thus, the sum in Equation (2.1) is $|\Omega_j|(1/N)(1/|\Omega_j|) = 1/N$. The conclusion is that Equation (2.1) holds for all colourings j, and we have that $\pi_{\mathcal{B}}$ is a stationary distribution. When $\mathcal{M}_{\text{scan}}$ is ergodic it follows that $\pi_{\mathcal{B}}$ is the unique stationary distribution. See [Ros06] for details on Markov chains and stationary distributions. □

2.3.3 Approximate counting and approximating the partition function

Exact counting of combinatorial structures is rarely possible in polynomial time. As demonstrated by Jerrum in [Jer03], the number of spanning trees in a graph, and the number of perfect matchings in a planar graph, are examples of structures that can be counted in polynomial time. However, counting in general is a hard problem, and counting the number of proper q-colourings of a graph is an example of this fact.

Counting problems can be viewed as functions $f : \Sigma^* \to \mathbb{N}$ mapping (encodings of) problem instances to natural numbers. An encoding x of an instance of a problem is a string of symbols from a finite alphapet Σ, and the length $|x|$ of the string x is proportional to the size of the instance. For example, if the problem is "How many proper 5-colourings does a given graph have?", then an encoding x of an instance to this problem represents a graph. The length of x is proportional to the size of the graph.

The analogue of the complexity class NP for counting problems was introduced

by Valiant [Val79] and is called #P. For an introduction to the complexity class NP, Turing machines and basic complexity theory, see for instance the books [GJ90, Pap94]. A counting problem $f : \Sigma^* \to \mathbb{N}$ belongs to the complexity class #P if there is a non-deterministic Turing machine M such that the number of accepting paths of M is $f(x)$ when given input x. A counting problem $f : \Sigma^* \to \mathbb{N}$ belongs to the complexity class FP if $f(x)$ can be computed in time polynomial in $|x|$. The complexity class FP $\subseteq$ #P. There are counting problems that are complete for the class #P [Jer03]. If there is an algorithm that solves a #P-complete problem in polynomial time then all counting problems in #P are solvable in polynomial time and hence FP = #P. Thus, as in the case of decision problems, it is possible to provide persuasive evidence for intractability of a counting problem if it can be shown to be #P-complete (or #P-hard, using terminology from standard complexity theory). The problem of counting the number of proper q-colourings of a graph with maximum degree Δ is #P-complete whenever $q \geq 3$ and $\Delta \geq 3$ [BDGJ99].

If exact counting is not possible (or known to be possible) in polynomial time, it does not rule out the possibility of *approximate counting*. If there is a known efficient method for *uniformly sampling* combinatorial structures, then this method can sometimes be utilised in order to approximately count the number of structures in an efficient way. In the case of proper colourings of a graph, approximate counting with good accuracy can be achieved by collecting sufficiently many samples of colourings using different boundary colourings. For details on the topic of how sampling and counting are related, see Jerrum, Valiant and Vazirani [JVV86]. For a concrete example of how to approximately count the number of matchings in a graph, given an efficient way of sampling matchings, see Jerrum [Jer03]. Note that the problem of exactly counting the number of matchings in a graph is #P-complete [Val79].

A counting problem can be generalised to a function $f : \Sigma^* \to \mathbb{R}$, where the range is not only natural numbers but real numbers. For example, for a fixed graph $\mathcal{G}$, q colours and a fixed value of λ in the anti-ferromagnetic Potts model, f might be the partition function (see Section 2.1.2). Input to f is (an encoding of) a region R and a q_0-colouring $\mathcal{B}$ of ∂R. Thus, if $\lambda = 0$ then f returns the number of proper q-colourings of R that agree with the boundary colouring $\mathcal{B}$.

A *randomised approximation scheme* for $f : \Sigma^* \to \mathbb{R}$ is a randomised algorithm that takes as input an instance $x \in \Sigma^*$ and an error tolerance $\varepsilon > 0$, and outputs a number $Y \in \mathbb{R}$. Here Y is a random variable, computed with some randomisation process in the algorithm, such that

$$\Pr[(1-\varepsilon)f(x) \leq Y \leq (1+\varepsilon)f(x)] \geq \frac{3}{4}.$$

The number $\frac{3}{4}$ above could be replaced by any number in the interval $(\frac{1}{2}, 1]$. By running

the algorithm sufficiently many times, the median y of all outputs Y will be between $(1-\varepsilon)f(x)$ and $(1+\varepsilon)f(x)$ with an arbitrary high probability. The randomised approximation scheme is said to be a *fully polynomial randomised approximation scheme*, or *FPRAS*, if it runs in time bounded by a polynomial in $|x|$ and $1/\varepsilon$.

Suppose that $f : \Sigma^* \to \mathbb{R}$ is the partition function under the anti-ferromagnetic Potts model for some graph $\mathcal{G}$ and parameters q and λ. Using ideas of Jerrum, Valiant and Vazirani [JVV86], an efficient sampling algorithm for $\pi_{\mathcal{B}}$ can be turned into an FPRAS for the partition function f. A straightforward proof is based on Dyer and Greenhill's extension [DG99] of [JVV86], however this is outside the scope of this thesis.

An example of an efficient algorithm for sampling colourings from $\pi_{\mathcal{B}}$ is a rapidly mixing Glauber dynamics. Recall that a rapidly mixing Glauber dynamics runs in time bounded by a polynomial in $|x|$ (the size of the graph) and $\log(1/\delta)$, where δ is the sampling tolerance. Note that the dependence of the run-time on the tolerance ε and δ is different for the FPRAS and the sampler: polynomial in $1/\varepsilon$ and $\log(1/\delta)$, respectively. As pointed out in [Jer03], this difference is deliberate. If the sampling tolerance δ is improved "drastically" (made very small) then the running time of the sampler will increase only slightly. This however, is not true for the error tolerance ε for the FPRAS.

2.4 Graphs of sub-exponential growth

All graphs $\mathcal{G}$ that we consider in this thesis have a property known as *uniformly sub-exponential*. The following definition is taken from [GMP05].

Definition 2.18 (Uniformly sub-exponential). *For a vertex v in $V_{\mathcal{G}}$, let $n_d(v)$ denote the number of vertices that are at distance d from v. (Note that $n_0(v) = 1$.) We say that $\mathcal{G}$ is* uniformly sub-exponential *if there exist a function $\kappa(d)$ such that*

(1) for all $a > 1$, $\kappa(d) \in o(a^d)$, and

(2) there exist $b_2 \geq b_1 > 0$ such that for all $v \in V_{\mathcal{G}}$, $b_1\kappa(d) \leq n_d(v) \leq b_2\kappa(d)$.

For any vertex $v \in V_{\mathcal{G}}$ and a non-negative integer d, let $\mathrm{Ball}_d(v)$ denote the set of vertices that are at most distance d from v. Thus we have $\mathrm{Ball}_0(v) = \{v\}$. The following definition of *neighbourhood-amenability* is taken from [GMP05].

Definition 2.19 (Neighbourhood-amenability). *For a non-negative integer d, let*

$$T_d = \sup_{v \in V_{\mathcal{G}}} \frac{|\partial \mathrm{Ball}_d(v)|}{|\mathrm{Ball}_d(v)|}.$$

The graph $\mathcal{G}$ is said to be neighbourhood-amenable *if* $\inf_d T_d = 0$.

We have the following lemma.

Lemma 2.20 ([GMP05], Lemma 39). *If the graph $\mathcal{G}$ is uniformly sub-exponential then $\mathcal{G}$ is neighbourhood-amenable.*

The implication from strong spatial mixing to rapidly mixing Glauber dynamics is only known to hold for graphs of sub-exponential growth [Wei06]. We will show that the existence of a K-coupling cover implies that the Glauber dynamics is rapidly mixing (Theorem 2.15). In order to do so we will use the fact that the graph is neighbourhood-amenable.

From Definition 2.19 above it follows that for a neighbourhood-amenable graph and any real number $a > 0$, there is an integer $d \geq 0$ such that

$$\frac{|\partial \mathrm{Ball}_d(v)|}{|\mathrm{Ball}_d(v)|} \leq a,$$

uniformly in v. Hence the "surface-area-to-volume" ratio can be made arbitrarily small with a suitable choice of radius d. Most natural lattices are neighbourhood-amenable. In this thesis we present proofs of strong spatial mixing and rapid mixing for the *square lattice* $\mathbb{Z}^2$ (also known as the *grid*), the *triangular lattice* and the *kagome lattice* (see Figure 2.3). These lattices are formally defined in the next sections and we prove that they are uniformly sub-exponential (and hence neighbourhood-amenable from Lemma 2.20).

2.4.1 The square lattice

The square lattice (Figure 2.3(a)), also denoted $\mathbb{Z}^2$, or the grid, is the infinite graph $\mathcal{G}$ with vertex set $V_{\mathcal{G}} = \{(x, y)\}$, where $x, y \in \mathbb{Z}$. The edge set $E_{\mathcal{G}}$ contains all edges between vertices that differ by one in exactly one coordinate. Thus, $\mathbb{Z}^2$ is a 4-regular graph and the maximum degree $\Delta = 4$.

Lemma 2.21. *The square lattice is uniformly sub-exponential.*

Proof. Let $n_d(v)$ be the number of vertices at distance d from a vertex v of the square lattice. We will show that $n_d(v) = 4d$ and then from Definition 2.18 it follows that $\mathbb{Z}^2$ is uniformly sub-exponential. Assume without loss of generality that v is the point of Cartesian coordinates $(0, 0)$. Let V_d denote the set of vertices at distance d from v.

For $x > 0$ and $y > 0$, there are exactly $d - 1$ vertices (x, y) in V_d, namely the vertices $(x, d - x)$, where $0 < x < d$. To see this, first suppose that $0 < x < d$. There is a shortest path of length d from v to $(x, d - x)$ that consists of x moves to the right followed by $d - x$ moves upwards. Thus a vertex (x, y), where $0 < y < d - x$ will not be in V_d. A vertex (x, y), where $y > d - x$ cannot belong to V_d either since following a shortest path from v to (x, y) would require more than d steps. Now suppose that

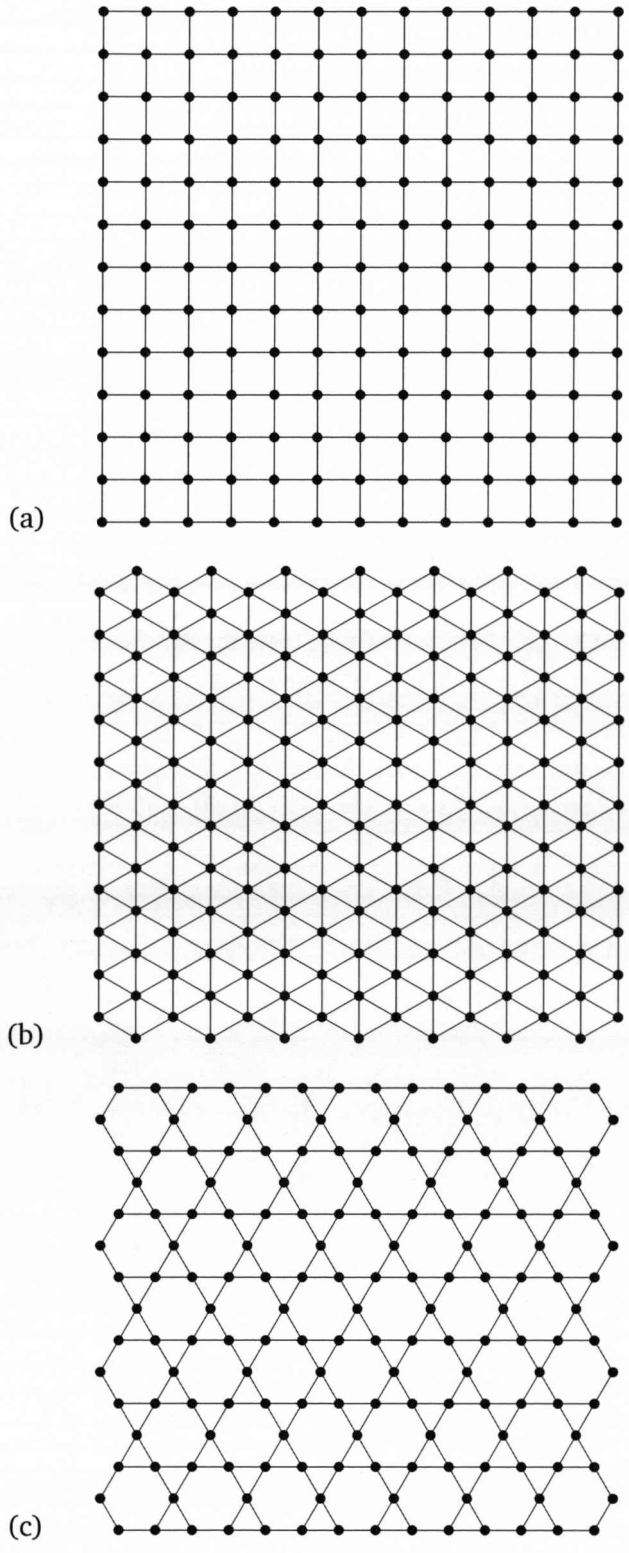
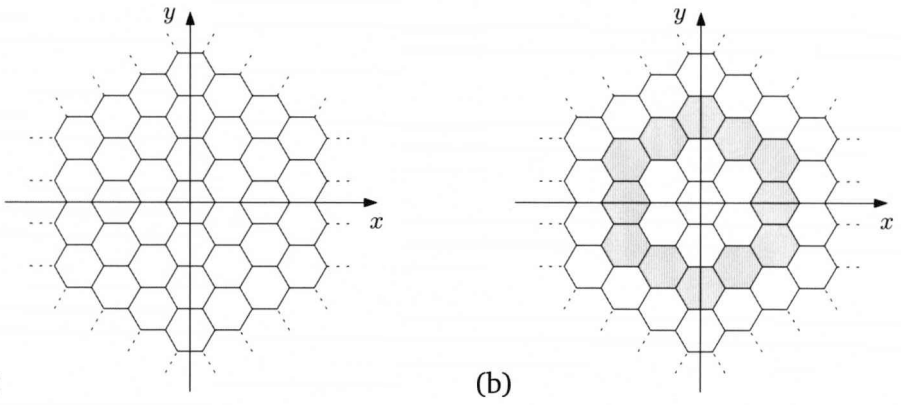


Figure 2.3. (a) The square lattice. (b) The triangular lattice. (c) The kagome lattice. The lattices are drawn in the traditional way in this figure, where a solid circle denotes a vertex.



(a) (b)

Figure 2.4. The triangular lattice, here drawn in a coordinate system. (b) Vertices at distance 2 from the origin are shaded.

$x \geq d$. If $x = d$ the a vertex (x, y) belongs to V_d only if $y = 0$. If $x > d$ the there is no vertex $(x, y) \in V_d$.

By symmetry we have that there are also $d-1$ vertices (x, y) in V_d for $x > 0$ and $y < 0$, $x < 0$ and $y > 0$, and $x < 0$ and $y < 0$, respectively. Thus, for $x \neq 0$ and $y \neq 0$ there are exactly $4d - 4$ vertices (x, y) in V_d. Now note that the vertices $(0, d)$, $(d, 0)$, $(0, -d)$ and $(-d, 0)$ are also in V_d. It is obvious that these are the only four vertices in V_d for which either $x = 0$ or $y = 0$. We have that $n_d(v) = 4d$, and from Definition 2.18 it follows that $\mathbb{Z}^2$ is uniformly sub-exponential. □

In Chapter 10 we are concerned with a systematic-scan approach for sampling colourings of a finite rectangular region of the square lattice. It will be convenient to work on the region as a *torus* in order to avoid technicalities regarding the vertices on the boundary of the region. A torus is formed from a rectangular region R of the square lattice by connecting the topmost vertices of R to the bottommost vertices of R, and by connecting the leftmost vertices of R to the rightmost vertices of R. Each vertex in the torus has degree four and the torus has no boundary vertices.

An $n \times m$-*block* is a rectangular region of the square lattice that has height n vertices and width m vertices.

2.4.2 The triangular lattice

The triangular lattice (Figure 2.3(b)) is the infinite graph $\mathcal{G}$ with vertex set $V_{\mathcal{G}} = \{(x, y)\}$, where $x, y \in \mathbb{Z}$ are both even or both odd. The lattice is 6-regular ($\Delta = 6$) and a vertex $(x, y) \in V_{\mathcal{G}}$ has the six neighbours $(x, y + 2)$, $(x + 1, y + 1)$, $(x + 1, y - 1)$, $(x, y - 2)$, $(x - 1, y - 1)$ and $(x - 1, y + 1)$ (see Figure 2.4(a), where faces represent vertices of the lattice).

Lemma 2.22. *The triangular lattice is uniformly sub-exponential.*

Proof. Let $n_d(v)$ be the number of vertices at distance d from a vertex v of the triangular lattice. We will show that $n_d(v) = 6d$ and then from Definition 2.18 it follows that the triangular lattice is uniformly sub-exponential. Assume without loss of generality that v is the point of Cartesian coordinates $(0, 0)$. Let V_d denote the set of vertices at distance d from v. Let U_d be the set of all vertices with Cartesian coordinates of the following form:

$$\begin{aligned}
(\ \ d, d-2x) &\qquad 0 < x < d \\
(-d, d-2x) &\qquad 0 < x < d \\
(\ \ x, 2d-x) &\qquad 0 \leq x \leq d \\
(-x, 2d-x) &\qquad 0 < x \leq d \\
(\ \ x, -2d+x) &\qquad 0 \leq x \leq d \\
(-x, -2d+x) &\qquad 0 < x \leq d
\end{aligned}$$

Note that the total number of vertices in U_d add up to $6d$. Also note that the sets $U_0, U_1, U_2, \ldots$ are all disjoint and the union of them is the whole vertex set of the triangular lattice. See Figure 2.4(b) for an example of vertices in U_2.

We will use induction on d to prove that $V_d = U_d$. The base case $d = 1$ is trivial. For the inductive step, assume that $V_d = U_d$ for $i = 1, \ldots, d-1$. In order to prove that $V_d = U_d$, we will prove two claims:

Claim 1. Every vertex $v \in U_d$ has a neighbour in U_{d-1}.

Claim 2. A vertex $v \in U_{d-1}$ has no neighbour in $U_{d'}$, where $d' > d$.

First we prove Claim 1. Suppose that $(d, d-2x)$, $0 < x < d$, is a vertex in U_d (the proof for vertices of the other five types is similar). The vertex $(d, d-2x)$ has six neighbours: $(d, d-2x+2)$, $(d+1, d-2x+1)$, $(d+1, d-2x-1)$, $(d, d-2x-2)$, $(d-1, d-2x-1)$, $(d-1, d-2x+1)$. We see that both the neighbours $(d-1, d-2x-1)$ and $(d-1, d-2x+1)$ are in U_{d-1}.

Now we prove Claim 2. Suppose that $(d-1, d-1-2x)$, $0 < x < d-1$, is a vertex in U_{d-1} (the proof for vertices of the other five types is similar). The vertex $(d-1, d-1-2x)$ has six neighbours: $(d-1, d-1-2x+2)$, $(d-1+1, d-1-2x+1)$, $(d-1+1, d-1-2x-1)$, $(d-1, d-1-2x-2)$, $(d-1-1, d-1-2x-1)$, $(d-1-1, d-1-2x+1)$. We see that every neighbour is in either U_{d-2}, U_{d-1} or U_d.

Both claims are proved and hence $V_d = U_d$. We have $n_d(v) = |V_d| = |U_d| = 6d$. From Definition 2.18 it follows that the triangular lattice is uniformly sub-exponential. □

2.4.3 The kagome lattice

The kagome lattice (Figure 2.3(c)) is a natural lattice of interest in statistical physics

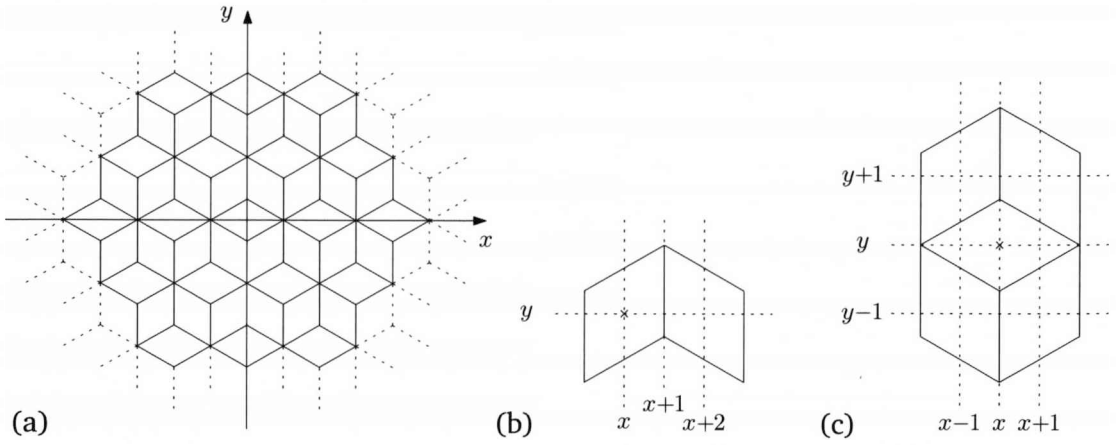


Figure 2.5. (a) The kagome lattice, here drawn in a coordinate system. (b) A vertex $(x, y) \in V_{\text{odd}}$ and its right neighbour. (c) A vertex $(x, y) \in V_{\text{even}}$ and its four neighbours.

(see [SS97]). The name kagome is derived from the Japanese word kagome, meaning the pattern of holes ("me", literally "eyes") in a basket ("kago"), as the lattice appears in woven baskets.

Let V_{odd} and V_{even} be the following two sets of Cartesian coordinates.

$$\begin{aligned} V_{\text{odd}} &= \{(x, y) \mid x, y \in \mathbb{Z} \text{ are both odd}\}, \\ V_{\text{even}} &= \{(x, y) \mid x = 4k_1 + r,\ y = 4k_2 + r,\ k_1, k_2 \in \mathbb{Z},\ r \in \{0, 2\}\}. \end{aligned}$$

The kagome lattice is the infinite graph $\mathcal{G}$ with vertex set $V_{\mathcal{G}} = V_{\text{odd}} \cup V_{\text{even}}$. Figure 2.5(a) illustrates the kagome lattice drawn in a coordinate system. Here faces represent vertices of the lattice. The two vertices in Figure 2.5(b) are in V_{odd}, and the middle vertex in Figure 2.5(c) is in V_{even}. Figure 2.5(b) shows that two adjacent vertices in V_{odd} differ by 2 in their x-coordinate. Figure 2.5(c) shows the position of the four neighbours of a vertex $(x, y) \in V_{\text{even}}$. Formally, the edge set $E_{\mathcal{G}}$ of the kagome lattice is defined as

$$\begin{aligned} E_{\mathcal{G}} = {} & \{((x, y), (x + 2, y)) \mid (x, y) \in V_{\text{odd}}\} \cup \\ & \{((x, y), (x + 1, y + 1)) \mid (x, y) \in V_{\text{even}}\} \cup \\ & \{((x, y), (x + 1, y - 1)) \mid (x, y) \in V_{\text{even}}\} \cup \\ & \{((x, y), (x - 1, y + 1)) \mid (x, y) \in V_{\text{even}}\} \cup \\ & \{((x, y), (x - 1, y - 1)) \mid (x, y) \in V_{\text{even}}\}. \end{aligned}$$

Lemma 2.23. *The kagome lattice is uniformly sub-exponential.*

Proof. Let $n_d(v)$ denote the number of vertices at distance d from a vertex v. We will show that $n_d(v) \in \Theta(d)$ for all vertices v. Then it follows from Definition 2.18 that the

triangular lattice is uniformly sub-exponential.

First assume that $v \in V_{\text{odd}}$. In order to derive lower and upper bounds on $n_d(v)$, we assume without loss of generality that $v = (1,1)$ is the vertex at x-coordinate 1 and y-coordinate 1. Fix any positive integer d.

We first derive a lower bound on $n_d((1,1))$. For each odd value of $y \in \{1,\ldots,d\}$, let (x,y) be the vertex at distance d from $(1,1)$ that is reached with the following path: $(1,1),(2,2),(3,3),\ldots,(y,y),(y+2,y),(y+4,y)\ldots,(x,y)$. Note that vertex $(y,y) \in V_{\text{odd}}$, and from (y,y) we go as far as possible to the right. Also note that there is no path from $(1,1)$ to (x,y) that is shorter than length d. Thus, there are at least $\lfloor d/2 \rfloor$ vertices at distance d from $(1,1)$, and we have $n_d((1,1)) \geq \lfloor d/2 \rfloor$.

When deriving an upper bound on $n_d((1,1))$ we will use two claims:

Claim 1. For any two vertices (x, y_{low}) and (x, y_{high}), where $1 \leq y_{\text{low}} < y_{\text{high}}$, the distance between $(1,1)$ and (x, y_{low}) is strictly smaller than the distance between $(1,1)$ and (x, y_{high}). We prove the claim by considering two cases:

Case (i). Assume that x is odd, and hence both (x, y_{low}) and (x, y_{high}) are in V_{odd}. Consider a shortest path from $(1,1)$ to (x, y_{high}). The path must use a vertex $(x_{\text{pass}}, y_{\text{low}}) \in V_{\text{odd}}$ at y-coordinate y_{low}. From $(x_{\text{pass}}, y_{\text{low}})$ we can reach (x, y_{low}) in exactly $|x - x_{\text{pass}}|/2$ steps. The number of steps required to reach (x, y_{high}) from $(x_{\text{pass}}, y_{\text{low}})$ is strictly greater than $|x - x_{\text{pass}}|/2$ since some steps must be used to increase the y-coordinate so it will eventually reach y_{high}, and for each such up-move the x-coordinate is increased/decreased only by 1. Thus, if x is odd then the distance between $(1,1)$ and (x, y_{low}) is strictly smaller than the distance between (x, y_{high}).

Case (ii). Assume that x is even, and hence both (x, y_{low}) and (x, y_{high}) are in V_{even}. We will use the same argument as for odd values of x, only with the difference that we consider a vertex $(x_{\text{pass}}, y_{\text{low}} - 1) \in V_{\text{odd}}$ on a shortest path from $(1,1)$ to (x, y_{high}). From $(x_{\text{pass}}, y_{\text{low}} - 1)$ we can reach (x, y_{low}) in at most $\lfloor |x - x_{\text{pass}}|/2 \rfloor + 1$ steps, where the $+1$ comes from the fact that we need to go up one y-coordinate. The number of steps required to reach (x, y_{high}) from $(x_{\text{pass}}, y_{\text{low}} - 1)$ is strictly greater than $\lfloor |x - x_{\text{pass}}|/2 \rfloor + 1$ since some steps must be used to increase the y-coordinate so it will eventually reach y_{high}, and for each such up-move the x-coordinate is increased/decreased only by 1. Thus, also for even values of x we have that the distance between $(1,1)$ and (x, y_{low}) is strictly smaller than the distance between $(1,1)$ and (x, y_{high}).

Claim 2. For any two vertices (x, y_{low}) and (x, y_{high}), where $y_{\text{low}} < y_{\text{high}} \leq 1$, the distance between $(1,1)$ and (x, y_{high}) is strictly smaller than the distance between $(1,1)$ and (x, y_{low}). We prove the claim by using exactly the same reasoning as for Claim 1.

Using Claim 1 and 2 we conclude that there are at most two vertices (x,y) and (x,y'), with the same x-coordinate, at distance d from $(1,1)$. The leftmost vertex that

is at distance d from from $(1,1)$ is $(1-2d,1)$. It is reached by making d consecutive left-moves. Similarly, the rightmost vertex at distance d from $(1,1)$ is $(1+2d,1)$. Thus, the x-coordinate of any vertex at distance d from $(1,1)$ is in the set $\{1-2d,\ldots,1+2d\}$, and hence there are at most $2 \times (4d+1) = 8d+2$ vertices at distance d from $(1,1)$. That is, $n_d((1,1)) \leq 8d+2$. We have now showed that for any vertex $v \in V_{\text{odd}}$, $\lfloor d/2 \rfloor \leq n_d(v) \leq 8d+2$.

It remains to derive upper and lower bounds on $n_d(v)$ for $v \in V_{\text{even}}$. Without loss of generality we assume that $v = (0,0)$ is the vertex at x-coordinate 0 and y-coordinate 0. Fix any positive integer d.

We derive a lower bound on $n_d((0,0))$ in the same way as when $v = (1,1)$. For each odd value of $y \in \{1,\ldots,d\}$, let (x,y) be the vertex at distance d from $(0,0)$ that is reached with the following path: $(0,0),(1,1),(2,2),\ldots,(y,y),(y+2,y),(y+4,y)\ldots,(x,y)$. Thus, there are at least $\lfloor d/2 \rfloor$ vertices at distance d from $(0,0)$, and we have $n_d((0,0)) \geq \lfloor d/2 \rfloor$.

We now derive an upper bound on $n_d((0,0))$. Vertex $(0,0)$ has exactly four neighbours: $(1,1)$, $(1,-1)$, $(-1,-1)$ and $(-1,1)$, which are all in V_{odd}. The shortest path from $(0,0)$ to any vertex at distance d from $(0,0)$ must use one of these four vertices. Thus, an upper bound on the number of vertices at distance d from $(0,0)$ is $n_d((0,0)) \leq n_{d-1}((1,1)) + n_{d-1}((1,-1)) + n_{d-1}((-1,-1)) + n_{d-1}((-1,1))$. From the upper bound above we have that there are at most $8(d-1)+2$ vertices at distance $d-1$ from a vertex in V_{odd}. Hence there are at most than $4 \times (8(d-1)+2) = 32d-24$ vertices at distance d from $(0,0)$, and we have $n_d((0,0)) \leq 32d-24$.

Finally, for any vertex v and any positive integer d, we have $\lfloor d/2 \rfloor \leq n_d(v) \leq 32d-24$. From Definition 2.18 it follows that the kagome lattice is uniformly sub-exponential. □

Chapter 3

Statement of results

The results presented and proved in this thesis are summarised in the theorems below. Theorem 3.1 is proved by hand but the other theorems are proved with computer assistance. Computer-assisted proofs of mixing, tailored for specific spin systems, can often provide better mixing bounds than hand proofs. In Section 3.1 we cover some previous results which can be compared to the improvements presented in this thesis.

Theorem 3.1. *Consider the anti-ferromagnetic Potts model on an infinite and uniformly sub-exponential graph $\mathcal{G}$ with maximum degree Δ and parameters q and $\lambda \in [0,1]$. Let $a = \lfloor(\Delta-1)/(q-2)\rfloor$, $r = (\Delta-1) \bmod (q-2)$ (so $a(q-2)+r = \Delta-1$) and*

$$\kappa = \frac{1-\lambda}{1+\lambda+r\lambda^{a+1}+(q-2-r)\lambda^{a}}.$$

The system has strong spatial mixing if $\kappa < 1/(\Delta-1)$. Furthermore, if $\kappa < 1/(\Delta-1)$, and $q \geq \Delta+2$ or $\lambda > 0$, then the Glauber dynamics is rapidly mixing for any region R and any q_0-colouring $\mathcal{B}$ of ∂R. The mixing time $\tau(\delta) \in O(n^2 + n\log\frac{1}{\delta})$, where n is the number of vertices in R.

Note, if $q \geq \Delta+1$ and $q > (2\Delta-1)(1-\lambda)$ then $\kappa < 1/(\Delta-1)$.

Theorem 3.2. *The anti-ferromagnetic Potts model on $\mathbb{Z}^2$ with parameters q and $\lambda \leq 1$, has strong spatial mixing in the following cases.*

(i) $q \geq 6$ *and* $\lambda \geq 0$,

(ii) $q = 5$ *and* $\lambda \geq 0.127$,

(iii) $q = 4$ *and* $\lambda \geq 0.262$, *and*

(iv) $q = 3$ *and* $\lambda \geq 0.393$.

Furthermore, for these values of q and λ, and any region R of $\mathbb{Z}^2$, the Glauber dynamics is rapidly mixing for any q_0-colouring $\mathcal{B}$ of ∂R. The mixing time $\tau(\delta) \in O(n^2 + n\log\frac{1}{\delta})$, where n is the number of vertices in R.

Theorem 3.3. *The anti-ferromagnetic Potts model on the triangular lattice with parameters $q = 9$ and $\lambda = 0$ has strong spatial mixing, and the Glauber dynamics is rapidly mixing for any region R and any q_0-colouring $\mathcal{B}$ of ∂R. The mixing time $\tau(\delta) \in O(n^2 + n \log \frac{1}{\delta})$, where n is the number of vertices in R.*

Theorem 3.4. *The anti-ferromagnetic Potts model on the kagome lattice with parameters $q = 5$ and $\lambda = 0$ has strong spatial mixing.*

Theorem 3.5. *For any region R of the kagome lattice, $q = 5$ and $\lambda = 0$, the Glauber dynamics is rapidly mixing for the 0-colouring $\mathcal{B}$ of ∂R. The mixing time $\tau(\delta) \in O(n^2 + n \log \frac{1}{\delta})$, where n is the number of vertices in R.*

Theorem 3.6. *Let R be a rectangular region of $\mathbb{Z}^2$ that is folded into a torus (hence $\partial R = \emptyset$). For $q = 7$ and $\lambda = 0$, the systematic scan Markov chain $\mathcal{M}_{\text{scan}}$ is rapidly mixing with 2×2-blocks. More precisely, the mixing time $\tau_{\mathcal{M}_{\text{scan}}}(\delta) \in O(\log \frac{n}{\delta})$, where n is the number of vertices in R.*

3.1 Overview of previous results

Salas and Sokal [SS97] have shown that a spin system under the anti-ferromagnetic Potts model has strong spatial mixing if the number of colours $q > 2\Delta$ and $\lambda \in [0, 1]$, where Δ is the maximum degree of the graph $\mathcal{G}$. However, Jerrum [Jer95] has pointed out that this result can be refined to the following theorem.

Theorem 3.7. *The anti-ferromagnetic Potts model on any infinite graph $\mathcal{G}$ with maximum degree Δ, and parameter $\lambda \in [0, 1]$, has strong spatial mixing if $q > 2\Delta(1 - \lambda)$.*

Theorem 3.1 improves this result and should been seen as the best general result known today, when the whole interval of λ is considered. For $\lambda = 0$ there are better results that apply. The zero-temperature case is particularly interesting since it admits only proper colourings.

In physics, graphs of interest are often lattice graphs of small maximum degree. It is therefore interesting to study spin systems with small Δ and few colours q. A general result by Goldberg, Martin and Paterson [GMP05] is stated in the following theorem. Note that the proofs of Theorem 3.1, 3.7 and 3.8 are hand proofs.

Theorem 3.8 ([GMP05], Theorem 5). *The anti-ferromagnetic Potts model on any infinite, triangle-free graph $\mathcal{G}$ with maximum degree $\Delta \geq 3$, and parameter $\lambda = 0$, has strong spatial mixing if $q \geq \alpha\Delta - \gamma$, where α is the solution to $\alpha^\alpha = e$ (so $\alpha \approx 1.76322$) and $\gamma = \frac{4\alpha^3 - 6\alpha^2 - 3\alpha + 4}{2(\alpha^2 - 1)} \approx 0.47031$.*

3.1.1 The square lattice (Glauber dynamics)

From Theorem 3.7 it follows that there is strong spatial mixing on $\mathbb{Z}^2$ for $q \geq 8$ and $\lambda \in (0, 1]$, and from Theorem 3.8 strong spatial mixing follows for $q \geq 7$ and $\lambda = 0$. Salas and Sokal [SS97] had earlier shown strong spatial mixing on $\mathbb{Z}^2$ for $q \geq 7$ and $\lambda = 0$, however, their proof is machine assisted. The case $q = 7$ and $\lambda = 0$ is also implied by another computer-assisted result due to Bubley, Dyer Greenhill and Jerrum [BDGJ99] that applies to 4-regular triangle-free graphs. Another machine-assisted result by Achlioptas, Molloy, Moore and van Bussel [AMMV04] states strong spatial mixing on $\mathbb{Z}^2$ for $q = 6$ and $\lambda = 0$. For $q = 2$ (see [Mar99]) it is known that there is a critical point λ_c such that there is strong spatial mixing for $\lambda > \lambda_c$ but not for $\lambda < \lambda_c$. The value $\lambda_c = \sqrt{2} - 1 \approx 0.414$ (see [SS97]). It is believed [SS97] that there is strong spatial mixing for every $\lambda \in (0, 1]$ for $q = 3$ and for every $\lambda \in [0, 1]$ for $q > 3$. The following proposition summarises the results discussed above and makes precise use of Theorem 3.7 for $q \leq 7$.

Proposition 1. *The anti-ferromagnetic Potts model on $\mathbb{Z}^2$ with parameters q and $\lambda \leq 1$, has strong spatial mixing in the following cases.*

(i) $q \geq 8$ *and* $\lambda \geq 0$,

(ii) $q = 7$ *and* $\lambda = 0$ *or* $\lambda \geq \frac{1}{8} = 0.125$,

(iii) $q = 6$ *and* $\lambda = 0$ *or* $\lambda \geq \frac{2}{8} = 0.25$,

(iv) $q = 5$ *and* $\lambda \geq \frac{3}{8} = 0.375$,

(v) $q = 4$ *and* $\lambda \geq \frac{4}{8} = 0.5$,

(vi) $q = 3$ *and* $\lambda \geq \frac{5}{8} = 0.625$, *and*

(vii) $q = 2$ *and* $\lambda > \sqrt{2} - 1 \approx 0.414$.

The results in Proposition 1 are improved by applying Theorem 3.1 on the square lattice. The new, better mixing bounds are stated in the following theorem.

Theorem 3.9. *The anti-ferromagnetic Potts model on $\mathbb{Z}^2$ with parameters q and $\lambda \leq 1$, has strong spatial mixing in the following cases.*

(i) $q \geq 7$ *and* $\lambda \geq 0$,

(ii) $q = 6$ *and* $\lambda = 0$ *or* $\lambda \geq \frac{1}{7} \approx 0.1429$,

(iii) $q = 5$ *and* $\lambda \geq \frac{2}{7} \approx 0.2857$,

(iv) $q = 4$ *and* $\lambda \geq \frac{1}{2}(\sqrt{33} - 5) \approx 0.3723$, *and*

(v) $q = 3$ *and* $\lambda \geq \lambda_0 \approx 0.4735$, *where* λ_0 *is the real solution of* $\lambda^3 + 4\lambda - 2 = 0$.

Furthermore, for these values of q *and* λ, *and any region* R *of* $\mathbb{Z}^2$, *the Glauber dynamics is rapidly mixing for any* q_0*-colouring* $\mathcal{B}$ *of* ∂R.

Theorem 3.2 improves these mixing bounds even further. Note that the proof of Theorem 3.2 is machine assisted, whereas the proof of Theorem 3.1 (and so Theorem 3.9) is a hand proof.

Finally it is worth pointing out that in the special case of $q = 3$ and $\lambda = 0$, two complementary results of Luby, Randall and Sinclair [LRS01] and Goldberg, Martin and Paterson [GMP04] give rapid mixing of the Glauber dynamics on a rectangular region of $\mathbb{Z}^2$ under the free boundary (0-colouring of the boundary).

3.1.2 The square lattice (systematic scan)

For the systematic scan Markov chain, rapid mixing on $\mathbb{Z}^2$ has been shown only for $q \geq 7$ and $\lambda = 0$ (Theorem 3.6). The proof of Theorem 3.6 is computer assisted. The previously best result known, due to Pedersen [Ped07] and Dyer, Goldberg and Jerrum [DGJ06a], states rapid mixing on $\mathbb{Z}^2$ for $q \geq 8$ and $\lambda = 0$. The result from Pedersen follows from a general hand-proved result; there is a systematic scan Markov chain that is rapidly mixing on general graphs with maximum degree Δ when $q \geq 2\Delta$, and each heat-bath update updates the colours of two adjacent vertices. More precisely, Pedersen shows that the systematic scan Markov chain mixes in $O(\log n)$ scans, where n is the number of vertices of the graph. Note that each scan updates $O(n)$ vertices. Earlier, Dyer, Goldberg and Jerrum [DGJ06a] had shown that there is a systematic scan Markov chain that mixes in $O(\log n)$ scans when $q > 2\Delta$, and in $O(n^2 \log n)$ scans when $q = 2\Delta$. In their version of the systematic scan Markov chain, only one vertex is updated in each heat-bath update.

3.1.3 The triangular lattice

Salas and Sokal provided in [SS97] a machine-assisted proof of strong spatial mixing on the triangular lattice for $q \geq 11$ and $\lambda = 0$. This was the best result known on the triangular lattice until Goldberg, Martin and Paterson [GMP05] applied the technique they used in their proof of Theorem 3.8 to produce a machine-assisted proof of strong spatial mixing for $q = 10$ and $\lambda = 0$. In this thesis, we refine their technique further and show (Theorem 3.3) strong spatial mixing on the triangular lattice for $q = 9$ and $\lambda = 0$. It is believed [SS97] that there is strong spatial mixing on the triangular lattice for $q \geq 5$ and $\lambda \in [0, 1]$. For more on conjectured mixing bounds for the triangular lattice, see [SS97].

3.1.4 The kagome lattice

It is believed [SS97] that there is strong spatial mixing on the kagome lattice for $q \geq 4$ and $\lambda \in [0, 1]$. From Theorem 3.7 it follows that there is strong spatial mixing on the kagome lattice for $q \geq 9$ and $\lambda = 0$, and $q \geq 8$ and $\lambda \in (0, 1]$. Since the kagome lattice contains triangles, Theorem 3.8 cannot be applied. Salas and Sokal provided in [SS97] a machine-assisted proof for $q \geq 6$ and $\lambda = 0$. In this thesis, Theorem 3.4 improves their result to $q \geq 5$ and $\lambda = 0$. With as few colours as five, the Glauber dynamics is not necessarily ergodic (states can be "frozen") if the colouring of the boundary of the region is chosen arbitrarily. Hence Theorem 3.5 states rapid mixing for five colours only when the boundary colouring is restricted to the free boundary (0-colouring).

Chapter 4

Rapidly mixing Glauber dynamics

In this chapter we prove Theorem 2.15. We do this similarly to how Dyer, Sinclair, Vigoda and Weitz proved that strong spatial mixing implies rapid mixing for $\mathbb{Z}^d$ in [DSVW04]. The proof is based on path coupling [BD97a, BD97b], which is a technique for deriving upper bounds on the mixing time of Markov chains. It will be convenient to consider a Markov chain, denoted $\mathcal{M}_d$, that is more general than the Glauber dynamics. The Markov chain $\mathcal{M}_d$ updates not necessarily only one vertex in each step, but updates a set of vertices that lie within a ball of radius $d \geq 0$. The Glauber dynamics is therefore the special case when the radius of the ball is $d = 0$. We will use path coupling to show that if the system has a K-coupling cover then there is an integer $d \geq 0$ such that $\mathcal{M}_d$ is rapidly mixing. Once the mixing time for $\mathcal{M}_d$ has been established we will use Markov chain comparison techniques to derive the mixing time of the Glauber dynamics $\mathcal{M}_0$.

The chapter is organised as follows. In Section 4.1 we define the Markov chain $\mathcal{M}_d$. In Section 4.2 we explain the concept of coupling, in particular path coupling. In Section 4.3 we use path coupling to derive the mixing time of $\mathcal{M}_d$. In Section 4.4 we use the mixing result of $\mathcal{M}_d$ to derive the mixing time of the Glauber dynamics.

4.1 The Markov chain $\mathcal{M}_d$

Recall from Section 2.4 that $\mathrm{Ball}_d(v)$ denotes the set of vertices that are at most distance d from a vertex $v \in V_\mathcal{G}$. Thus $\mathrm{Ball}_0(v) = \{v\}$. For any region R, let $R_v^d = R \cap \mathrm{Ball}_d(v)$ and

$$R^d = \{v \in V_\mathcal{G} \mid R_v^d \neq \emptyset\}.$$

For a fixed $d \geq 0$ we define the heat-bath Markov chain $\mathcal{M}_d$ as follows.

Definition 4.1 (Markov chain $\mathcal{M}_d$). *For any region R, any q_0-colouring $\mathcal{B}$ of ∂R and any integer $d \geq 0$, the* heat-bath Markov chain $\mathcal{M}_d$ *is defined as follows.*

- *The state space is Ω_R^+ if $\lambda > 0$ and $\Omega_R(\mathcal{B})$ if $\lambda = 0$.*

- *A transition is made from a state σ to σ' in the following way:*

  1. *Choose a vertex v uniformly at random from R^d.*
  2. *Obtain the new colouring σ' by performing a heat-bath update of σ with respect to R_v^d.*

It is easy to verify that the definition of $\mathcal{M}_0$ is equivalent to Definition 2.14 of the Glauber dynamics. For any d, the stationary distribution of $\mathcal{M}_d$ is $\pi_{\mathcal{B}}$, which is stated in the following lemma. Note that $\mathcal{M}_d$ is ergodic for all $d \geq 0$ if $\mathcal{M}_0$ is ergodic.

Lemma 4.2. *If $\mathcal{M}_d$ is ergodic then the stationary distribution is $\pi_{\mathcal{B}}$.*

Proof. Consider any two states σ and σ' of $\mathcal{M}_d$ such that there is a positive probability of making a transition from σ to σ'. Let $E_\sigma \subseteq E(R)$ be the set of edges in $E(R)$ that are monochromatic under σ and $\mathcal{B}$. Similarly, let $E_{\sigma'} \subseteq E(R)$ be the set of edges in $E(R)$ that are monochromatic under σ' and $\mathcal{B}$. Let $M = E_\sigma \cap E_{\sigma'}$ be the set of edges that are monochromatic under both σ and σ'. Let $M_\sigma = E_\sigma \backslash M$ be set of edges that are monochromatic under σ but not under σ'. Let $M_{\sigma'} = E_{\sigma'} \backslash M$ be set of edges that are monochromatic under σ' but not under σ. For each vertex $v \in R^d$, let $M_v = E(R_v^d) \cap M$ be the set of edges that are monochromatic under both σ and σ', and have at least one endpoint in R_v^d. Let $m = |M|$, $m_\sigma = |M_\sigma|$, $m_{\sigma'} = |M_{\sigma'}|$ and $m_v = |M_v|$. We have

$$\pi_{\mathcal{B}}(\sigma) = \frac{\lambda^{m+m_\sigma}}{Z},$$

where Z is the normalising factor such that

$$\sum_{\theta \in \Omega_{\mathcal{M}_d}} \pi_{\mathcal{B}}(\theta) = 1,$$

where $\Omega_{\mathcal{M}_d}$ is the state space of $\mathcal{M}_d$. Note that Z does not depend on σ or σ'. Let V be the set of vertices v such that R_v^d contains all vertices on which σ and σ' differ. The probability of making a transition from σ to σ' is

$$\Pr[\sigma, \sigma'] = \sum_{v \in V} \left(\frac{1}{|R^d|} \cdot \frac{\lambda^{m_v + m_{\sigma'}}}{Z_v} \right),$$

where $1/|R^d|$ is the probability of choosing a ball centred at v, and $\lambda^{m_v+m_{\sigma'}}/Z_v$ is the probability of obtaining the new colouring σ' by performing a heat-bath update of σ with respect to R_v^d. The value Z_v is a normalising factor and depends only on R_v^d and the colouring of ∂R_v^d. Note that the colouring of ∂R_v^d does not depend on the vertices

on which σ and σ' differ. Now,

$$\pi_{\mathcal{B}}(\sigma)\Pr[\sigma,\sigma'] = \frac{\lambda^{m+m_\sigma}}{Z}\sum_{v\in V}\left(\frac{1}{|R^d|}\cdot\frac{\lambda^{m_v+m_{\sigma'}}}{Z_v}\right) = \frac{1}{Z\cdot|R^d|}\sum_{v\in V}\frac{\lambda^{m+m_\sigma+m_v+m_{\sigma'}}}{Z_v},$$

By interchanging σ and σ' it follows that $\pi_{\mathcal{B}}(\sigma)\Pr[\sigma,\sigma'] = \pi_{\mathcal{B}}(\sigma')\Pr[\sigma',\sigma]$. Thus $\mathcal{M}_d$ is reversible with respect to $\pi_{\mathcal{B}}$, and $\pi_{\mathcal{B}}$ is therefore a stationary distribution of $\mathcal{M}_d$. If $\mathcal{M}_d$ is ergodic then $\pi_{\mathcal{B}}$ is the unique stationary distribution. □

4.2 Coupling

Let $\mathcal{M}$ be any Markov chain on state space Ω and with transition matrix $P_{\mathcal{M}}$. A *(Markovian) coupling* based on $\mathcal{M}$ is a stochastic process (X_t, Y_t) on state space $\Omega\times\Omega$ such that each of (X_t) and (Y_t), considered marginally, is a faithful copy of $\mathcal{M}$. In other words, (X_t, Y_t) is a Markov chain on state space $\Omega\times\Omega$ with transition probabilities defined by

$$\Pr[X_1 = x' \mid X_0 = x,\ Y_0 = y] = P_{\mathcal{M}}(x,x'),$$
$$\Pr[Y_1 = y' \mid X_0 = x,\ Y_0 = y] = P_{\mathcal{M}}(y,y').$$

A cartoonish way of thinking of coupling is to pretend it represents the movements of a cat (X_t) and a mouse (Y_t), both jumping from state to state in the Markov chain $\mathcal{M}$. They both make their moves according to the transition probabilities of $\mathcal{M}$. Their movements could be independent, however, the idea of coupling is to specify some correlation that brings the cat together with the mouse in a small number of steps. That is, the cat is hunting the mouse. If there exist a coupling that brings the cat to the mouse in a sufficiently small number of steps, then it follows that $\mathcal{M}$ is rapidly mixing. This is formalised in the *coupling lemma* below. For details and a proof of the lemma, see for example [Ald83, Dia88, Jer03].

Lemma 4.3 (Coupling lemma). *Let (X_t, Y_t) be any coupling based on a Markov chain $\mathcal{M}$ on state space Ω. Suppose $t : [0,1] \to \mathbb{Z}^+$ is a function satisfying the condition: for all $x, y \in \Omega$, and all $0 < \delta < 1$,*

$$\Pr[X_{t(\delta)} \neq Y_{t(\delta)} \mid X_0 = x,\ Y_0 = y] \leq \delta.$$

Then the mixing time $\tau_{\mathcal{M}}(\delta)$ of $\mathcal{M}$ is bounded above by $t(\delta)$.

Note that once (X_t) and (Y_t) coalesce, at say time $t = t_0$, we can arrange for X_t to be equal to Y_t for all $t > t_0$. That is, (X_t) and (Y_t) make the same transition in each step.

The task is to design a coupling based on $\mathcal{M}_d$ such that $t(\delta)$ in Lemma 4.3 is bounded by a polynomial in n and $\log(1/\delta)$, where n is the number of vertices in the

region R. Showing the existence of a coupling with this property is not always easy. One great difficulty is to show that the probability in Lemma 4.3 is bounded by δ for *all* pairs of states (x, y) in $\Omega \times \Omega$. A powerful method called *path coupling*, invented by Bubley and Dyer [BD97a, BD97b], is an engineering solution to this problem. When using the path-coupling technique it is not necessary to explicitly consider all pairs of states (x, y). The only pairs of states one need to consider are pairs of *adjacent* states. For the Markov chain $\mathcal{M}_d$, two states (colourings) σ and σ' are adjacent only if they differ on vertices in R_v^d, for some vertex $v \in R^d$.

4.2.1 Path coupling

Before formally stating the consequence of path coupling, we describe the method by using the cartoonish example of coupling from the previous section, in which a cat (X_t) is hunting a mouse (Y_t) in the Markov chain $\mathcal{M}$. By defining a suitable metric on the state space, each pair of states will be at a certain distance apart. The general idea is to show that regardless of which states the cat and the mouse initially start at, the distance between them will decrease exponentially in the number of moves they make. Hence the cat will reach the mouse in a small number of steps.

One step of the coupling (X_t, Y_t) can be thought of in the following way: First the mouse (Y_t) makes a move according to the transition probabilities of $\mathcal{M}$. Then the cat (X_t) makes a move according to the transition probabilities of $\mathcal{M}$. The cat does not move independently of the mouse but is trying to correlate its move with the mouse's move in a clever way in order to get closer. Here closer suggests that there is a defined metric on the state space. We pretend that the states are islands and some states are connected with bridges. Two states that are connected with a bridge are said to be adjacent. Assume that any two states are connected with a path of adjacent states. The distance between two states is defined to be the length of the shortest path between them. Hence the distance between any two adjacent states is one. Note that the location of the bridges (the metric) does not have anything to do with the transition probabilities of the Markov chain. The cat and the mouse jump from state to state without necessarily using the bridges (might sound unusual for a cat). Now, position the mouse at any state Y_t and position the cat at any state X_t. Let $w_0, \ldots, w_k$ be any shortest-length path between them, where $Y_t = w_0$ and $X_t = w_k$. Along the path we position some other mouse-hungry cats (Z_t^i). That is, cat (Z_t^i) is positioned at state $Z_t^i = w_i$, $i = 1, \ldots, k-1$. The coupling (X_t, Y_t) is now defined by composing the couplings (Z_t^i, Z_t^{i+1}) along the path. First, the mouse makes a move from Y_t to Y_{t+1}. Then cat (Z_t^1), who is at a state adjacent to where the mouse was at time t, makes a move from Z_t^1 to Z_{t+1}^1. Now, this move is correlated (in a clever way) with the mouse's move in order to catch the mouse. Then cat (Z_t^2) makes a move from Z_t^2 to Z_{t+1}^2 that is

correlated with the move of cat Z_t^1, and so on. This chain of moves propagates through all cats, with cat (X_t) making a move based on the move of cat (Z_t^{k-1}). In other words, each cat bases its move on the move of the cat that was (at least originally) closer to the mouse. The distance between any two adjacent states is one, and hence the distance between the mouse (Y_t) and the cat (X_t) is k (originally). If we can show that for two cats (Z_t^i) and (Z_t^{i+1}) (and for the mouse (Y_t) and the cat Z_t^1), starting at any two adjacent states, the expected distance between them after one move is at most $\varphi < 1$, then, using path coupling, the expected distance between the mouse (Y_t) and the cat (X_t) after one move is at most $\varphi k < k$. After the second move it follows that the expected distance between the cat and the mouse is at most $\varphi^2 k$. After the third move the expected distance is at most $\varphi^3 k$, and so on. Thus, the expected distance between the cat and the mouse is decreasing exponentially regardless of which two states they started at. This is exactly the desired property of the coupling (X_t, Y_t) we are after, since it implies that the probability $\Pr[X_t \neq Y_t]$ decreases quickly with the time t. For details and more on path coupling, see [BD97a, BD97b, DG99].

The strength of path coupling is that an upper bound on the mixing time can be obtained by analysing how the distance changes between (X_t) and (Y_t) on two adjacent states. The following theorem, adapted from [DG99], formally summarises the path-coupling method.

Theorem 4.4 (Path coupling, [BD97a, DG99]). *Let $\mathcal{M}$ be a Markov chain on state space Ω, and let S be a relation $S \subseteq \Omega^2$ such that S has transitive closure Ω^2. Let $\phi : S \to \mathbb{Z}^+$ be a "proximity function" defined on pairs in S. We use ϕ to define a function Φ on Ω^2 as follows: For each pair $(x, y) \in \Omega^2$, let*

$$\Phi(x, y) = \min_{w_0, \ldots, w_k} \sum_{i=0}^{k-1} \phi(w_i, w_{i+1}),$$

where the minimum is over all paths $x = w_0, \ldots, w_k = y$ such that, for all $i \in [0, k-1]$, $(w_i, w_{i+1}) \in S$. Let (X_t, Y_t) be a coupling based on $\mathcal{M}$ defined over all pairs in S. Suppose for this coupling there is a $\varphi < 1$ such that for all $(x, y) \in S$ we have

$$\mathbb{E}[\Phi(X_{t+1}, Y_{t+1}) \mid (X_t, Y_t) = (x, y)] \leq \varphi \Phi(x, y).$$

Let D be the maximum value that Φ achieves on Ω^2. Then the mixing time of $\mathcal{M}$ is

$$\tau_{\mathcal{M}}(\delta) \leq \frac{\ln \frac{D}{\delta}}{1 - \varphi}.$$

In order to apply path coupling to a particular problem, the relation S and the proximity function ϕ in Theorem 4.4 have to be defined. In particular, the function Φ has to be easily deducible from ϕ. The next section will demonstrate how path coupling

can be applied to prove that the Markov chain $\mathcal{M}_d$ is rapidly mixing for a suitable choice of d.

4.3 Rapidly mixing $\mathcal{M}_d$

We prove the following lemma.

Lemma 4.5. *Consider the anti-ferromagnetic Potts model on a uniformly sub-exponential graph $\mathcal{G}$ with maximum degree Δ, and parameters q and λ. If the spin system has a K-coupling cover then there is an integer $d \geq 0$ such that $\mathcal{M}_d$ is rapidly mixing on any region R and q_0-colouring $\mathcal{B}$ of ∂R. More precisely, the mixing time of $\mathcal{M}_d$ is $\tau_{\mathcal{M}_d}(\delta) \in O(n \log \frac{n}{\delta})$, where n is the number of vertices in R.*

Proof. We use path coupling to prove the lemma. When defining the metric on the state space of $\mathcal{M}_d$ it will help to consider all colourings in Ω_R^+ even when $\lambda = 0$. Therefore we introduce a Markov chain called $\mathcal{M}_d^+$ which is defined exactly like $\mathcal{M}_d$, only with the difference that the state space of $\mathcal{M}_d^+$ is Ω_R^+ for all values of λ. Once an upper bound on the mixing time of $\mathcal{M}_d^+$ has been established it follows that the same upper bound also holds for $\mathcal{M}_d$ since they behave identically on the state space of $\mathcal{M}_d$. Note that we always assume $q \geq \Delta + 1$ when $\lambda = 0$ so the heat-bath move of $\mathcal{M}_d^+$ always selects a colouring from a distribution with non-empty support.

Let the Markov chain $\mathcal{M}$ in Theorem 4.4 be $\mathcal{M}_d^+$ (where d is to be specified later) and let the relation S be the set of pairs of colourings in Ω_R^+ that differ at a single vertex in R. Let $\phi(\sigma_1, \sigma_2) = 1$ for each $(\sigma_1, \sigma_2) \in S$, and hence, for any $\sigma_1, \sigma_2 \in \Omega_R^+$, $\Phi(\sigma_1, \sigma_2)$ is the Hamming distance between σ_1 and σ_2. We describe a coupled transition from the pair $(\sigma_1, \sigma_2) \in S$ to a new pair of colourings (σ_1', σ_2') and show that if d is large enough, then the expected distance $\mathbb{E}[\Phi(\sigma_1', \sigma_2')] \leq \varphi$ for a fixed $\varphi < 1$.

Recall Definition 4.1 of $\mathcal{M}_d$ and let (X_t, Y_t) be the coupling based on $\mathcal{M}_d^+$ that we describe next. Let $X_t = \sigma_1$ and $Y_t = \sigma_2$, where σ_1 and σ_2 differ only on a vertex $w \in R$. In one step of the Markov chain (X_t, Y_t) we choose the same vertex v uniformly at random from R^d for both (X_t) and (Y_t). Hence the heat-bath update will update the vertices in R_v^d for both (X_t) and (Y_t). Let $\mathcal{B}_1$ be the colouring of ∂R_v^d induced by σ_1 and $\mathcal{B}$, and let $\mathcal{B}_2$ be the colouring of ∂R_v^d induced by σ_2 and $\mathcal{B}$. Let (σ_1', σ_2') be the state (X_t, Y_t) makes a transition to. There are now three scenarios that can happen:

(1) $w \notin (R_v^d \cup \partial R_v^d)$,

(2) $w \in R_v^d$, or

(3) $w \in \partial R_v^d$.

If scenario (1) happens, then we update R_v^d with the same colouring for both (X_t) and (Y_t), which is possible since $\mathcal{B}_1$ and $\mathcal{B}_2$ are identical. The colourings σ_1' and σ_2' will differ only at vertex w and the Hamming distance between them is therefore still 1.

If scenario (2) happens, then we also update R_v^d with the same colouring for both (X_t) and (Y_t). However, colouring σ_1' and σ_2' will now be identical since $w \in R_v^d$, and the Hamming distance between them decreases by 1. Scenario (2) happens if and only if v is any of the vertices in $\mathrm{Ball}_d(w)$, and hence the probability of decreasing the Hamming distance by 1 is $|\mathrm{Ball}_d(w)|/|R^d|$.

If scenario (3) happens, then $\mathcal{B}_1$ and $\mathcal{B}_2$ differ on vertex w. Since the colour of w is not 0 in either $\mathcal{B}_1$ and $\mathcal{B}_2$, and the spin system is assumed to have a K-coupling cover, there is a coupling Ψ of $\pi_{\mathcal{B}_1}$ and $\pi_{\mathcal{B}_2}$ such that the expected Hamming distance between two colourings drawn from Ψ is at most K. We update R_v^d in (X_t) and (Y_t) by choosing two colourings from Ψ, and hence the expected Hamming distance between σ_1' and σ_2' is guaranteed to increase by at most K. Scenario (3) can only happen if $v \in \partial\mathrm{Ball}_d(w)$. If w is adjacent to the boundary ∂R there could be vertices in $\partial\mathrm{Ball}_d(w)$ that do not belong to R_d. With $\mathcal{H}_d(w) = \partial\mathrm{Ball}_d(w) \cap R_d$, it follows that the probability of increasing the expected Hamming distance by K is $|\mathcal{H}_d(w)|/|R^d|$, which is upper-bounded by $|\partial\mathrm{Ball}_d(w)|/|R^d|$.

Adding it all up, we see that the expected Hamming distance $\mathbb{E}[\Phi(\sigma_1', \sigma_2')]$ between σ_1' and σ_2' after one step of (X_t, Y_t) is upper-bounded by

$$1 - 1 \cdot \frac{|\mathrm{Ball}_d(w)|}{|R^d|} + K \cdot \frac{|\partial\mathrm{Ball}_d(w)|}{|R^d|} = 1 - \frac{|\mathrm{Ball}_d(w)| - K|\partial\mathrm{Ball}_d(w)|}{|R^d|}. \tag{4.1}$$

The graph $\mathcal{G}$ is uniformly sub-exponential and therefore also neighbourhood-amenable (Lemma 2.20). Hence there is a $d \geq 0$ such that

$$\frac{|\partial\mathrm{Ball}_d(w)|}{|\mathrm{Ball}_d(w)|} \leq \frac{1}{2K}, \tag{4.2}$$

uniformly in w. If we choose d such that $|\mathrm{Ball}_d(w)| \geq 2$ and Equation (4.2) holds, then $|\mathrm{Ball}_d(w)| - K|\partial\mathrm{Ball}_d(w)| \geq 1$ and the expected Hamming distance between σ_1' and σ_2' in Equation (4.1) is

$$\mathbb{E}[\Phi(\sigma_1', \sigma_2')] \leq 1 - \frac{1}{|R^d|}.$$

So for Theorem 4.4 we can take $\varphi = 1 - 1/|R^d|$, and the mixing time of $\mathcal{M}_d^+$ is

$$\tau_{\mathcal{M}_d^+}(\delta) \leq \frac{\ln \frac{D}{\delta}}{1 - \varphi} = |R^d| \ln \frac{D}{\delta}.$$

Since we use Hamming distance, we have that the maximum value D of Φ is n, where n is the number of vertices in R. Using the fact that d and the maximum degree Δ of $\mathcal{G}$

are constants, we see that $|R^d| \in O(n)$. This implies that

$$\tau_{\mathcal{M}_d^+}(\delta) \in O(n \log \frac{n}{\delta}).$$

Finally, $\mathcal{M}_d = \mathcal{M}_d^+$ when $\lambda > 0$. If $\lambda = 0$ then the state space of $\mathcal{M}_d$ is $\Omega_R(\mathcal{B})$. By starting the chain $\mathcal{M}_d^+$ on a proper colouring in $\Omega_R(\mathcal{B})$ we ensure that $\mathcal{M}_d^+$ behaves exactly like $\mathcal{M}_d$ and will never get to a colouring that is not in $\Omega_R(\mathcal{B})$. We therefore conclude that the upper bound on the mixing time $\tau_{\mathcal{M}_d^+}(\delta)$ also holds for $\mathcal{M}_d$. □

An upper bound on the mixing time of the Markov chain $\mathcal{M}_d$, for a fixed ball with radius $d \geq 0$, has now been established. In order to prove Theorem 2.15 we will use Lemma 4.5 to derive an upper bound on the mixing time of the Glauber dynamics $\mathcal{M}_0$. In the next section we use a Markov chain comparison technique to do so.

4.4 Rapidly mixing Glauber dynamics

We will compare the mixing time of the Markov chain $\mathcal{M}_d$ and $\mathcal{M}_0$ by using a method of Diaconis and Saloff-Coste [DSC93]. Their method has been used before by Goldberg, Martin and Paterson in [GMP05] to compare the mixing time of $\mathcal{M}_d$ and $\mathcal{M}_0$, where the Markov chains are defined only on the set of proper colourings ($\lambda = 0$). Next we review their comparison in [GMP05], but extend it to any value of $\lambda \in [0, 1]$. For a survey on Markov chain comparison in general, see [DGJM06].

The idea is to "simulate" the Markov chain $\mathcal{M}_d$ by running the chain $\mathcal{M}_0$, which only updates a single vertex in each step. If we know the mixing time of $\mathcal{M}_d$ and the "simulation" of $\mathcal{M}_d$ can be done efficiently, then it follows that the mixing time of $\mathcal{M}_0$ is not very different from the mixing time of $\mathcal{M}_d$. That is, if $\mathcal{M}_d$ is rapidly mixing then so is $\mathcal{M}_0$.

Let P_d and P_0 denote the transition matrix for the chain $\mathcal{M}_d$ and $\mathcal{M}_0$, respectively. For $i \in \{0, d\}$, let E_i be the set of pairs of distinct colourings (σ_1, σ_2) with $P_i(\sigma_1, \sigma_2) > 0$. The set E_i can be thought of as containing the edges of the transition graph of $\mathcal{M}_i$, and hence we sometimes refer to a pair in E_i as an edge. For every edge $(\sigma_1, \sigma_2) \in E_d$, let $\mathcal{P}_{\sigma_1,\sigma_2}$ be the set of paths from σ_1 to σ_2 using transitions of $\mathcal{M}_0$. More formally, let $\mathcal{P}_{\sigma_1,\sigma_2}$ be the set of paths $\gamma = (\sigma_1 = \theta_0, \theta_1, \ldots, \theta_k = \sigma_2)$ such that

(1) each (θ_i, θ_{i+1}) is in E_0, and

(2) each edge in E_0 appears at most once on γ.

We write $|\gamma|$ to denote the edge length of path γ. So, for example, if $\gamma = (\theta_0, \ldots, \theta_k)$ we have $|\gamma| = k$. Let $\mathcal{P} = \cup_{(\sigma_1,\sigma_2)\in E_d} \mathcal{P}_{\sigma_1,\sigma_2}$ be the set of all paths for all edges in E_d.

A *flow* is a function ϕ from $\mathcal{P}$ to the interval $[0,1]$ such that for every $(\sigma_1,\sigma_2) \in E_d$,

$$\sum_{\gamma \in \mathcal{P}_{\sigma_1,\sigma_2}} \phi(\gamma) = P_d(\sigma_1,\sigma_2)\pi_{\mathcal{B}}(\sigma_1).$$

For every $(\theta_1,\theta_2) \in E_0$, the *congestion of edge* (θ_1,θ_2) in the flow ϕ is the quantity

$$A_{\theta_1,\theta_2}(\phi) = \frac{1}{\pi_{\mathcal{B}}(\theta_1)P_0(\theta_1,\theta_2)} \sum_{\substack{\gamma \in \mathcal{P} \\ \text{such that} \\ (\theta_1,\theta_2)\in\gamma}} |\gamma|\phi(\gamma).$$

The *congestion of the flow* is the quantity

$$A(\phi) = \max_{(\theta_1,\theta_2)\in E_0} A_{\theta_1,\theta_2}(\phi).$$

Theorem 4.6 below describes how the mixing times of $\mathcal{M}_d$ and $\mathcal{M}_0$ are related. A proof of this theorem can be found in [DGJM06, Observation 13]. As pointed out in [GMP05], this theorem is similar to Proposition 4 of Randall and Tetali [RT00] except that the latter requires the eigenvalues of transition matrices to be non-negative. Both results are based closely on the ideas of Aldous [Ald83], Diaconis and Stroock [DS91], and Sinclair [Sin92].

Theorem 4.6. *Let $\Omega_{\mathcal{M}}$ be the state space of $\mathcal{M}_d$ and $\mathcal{M}_0$. Suppose that ϕ is a flow. Let $p = \min_{\theta\in\Omega_{\mathcal{M}}} P_0(\theta,\theta)$ and assume that $p > 0$. Then for any $0 < \delta' < \frac{1}{2}$*

$$\tau_{\mathcal{M}_0}(\delta) \le \ln\frac{1}{\delta\cdot\pi_{\min}} \cdot \max\left[A(\phi)\left(\frac{\tau_{\mathcal{M}_d}(\delta')}{\ln\frac{1}{2\delta'}}+1\right), \frac{1}{2p}\right]$$

where $\pi_{\min} = \min_{\sigma\in\Omega_{\mathcal{M}}} \pi_{\mathcal{B}}(\sigma)$.

We use Theorem 4.6 to prove the following lemma.

Lemma 4.7. *Suppose that the mixing time of $\mathcal{M}_d$ is $\tau_{\mathcal{M}_d}(\delta) \in O(n\log\frac{n}{\delta})$ (as in Lemma 4.5) and suppose that there is a flow ϕ such that the congestion $A(\phi) \in O(1)$. Then the mixing time of $\mathcal{M}_0$ (Glauber dynamics) is*

$$\tau_{\mathcal{M}_0}(\delta) \in O(n(n+\log\frac{1}{\delta})).$$

Proof. We apply Theorem 4.6. From Definition 2.14 of Glauber dynamics it follows that $p = \min_{\theta\in\Omega_{\mathcal{M}}} P_0(\theta,\theta) \ge \lambda^{\Delta}/q$. Let n be the number of vertices in R and suppose we take $\delta' = 1/n$. Then we have $\tau_{\mathcal{M}_d}(\delta') \in O(n\log n)$. Since $A(\phi) \in O(1)$, Theorem 4.6 gives

$$\tau_{\mathcal{M}_0}(\delta) \le \ln\frac{1}{\delta\cdot\pi_{\min}} \cdot O(1)\cdot O(n) = \left(\ln\frac{1}{\pi_{\min}} + O(\log\frac{1}{\delta})\right)\cdot O(n).$$

The number of monochromatic edges in a colouring of R is upper-bounded by $|E(R)| \leq n\Delta$, and the partition function $Z_{\mathcal{B}}$ is upper-bounded by the number of colourings $|\Omega_R^+| = q^n$ (see Section 2.1.2). Hence

$$\pi_{\min} \geq \frac{\lambda^{n\Delta}}{q^n}$$

and $\ln(1/\pi_{\min}) \in O(n)$, which makes $\tau_{\mathcal{M}_0}(\delta) \in O(n(n + \log \frac{1}{\delta}))$. □

In order to prove Theorem 2.15 it remains to construct a flow ϕ such that the congestion $A(\phi) \in O(1)$. The following lemma ensures that such a congestion can be constructed.

Lemma 4.8. *Consider any region R and any q_0-colouring $\mathcal{B}$ of ∂R. If $\lambda > 0$ or $q \geq \Delta + 2$ then there is a flow ϕ such that the congestion $A(\phi) \in O(1)$.*

Proof. Let $\prec$ be a fixed canonical ordering of the vertices in R. For every pair $(\sigma_1, \sigma_2) \in E_d$ we know that σ_1 and σ_2 differ only on vertices that are contained in $\mathrm{Ball}_d(v)$ for some vertex $v \in R^d$. Let V be the set of vertices on which they differ. Let $\gamma_{\sigma_1,\sigma_2} \in \mathcal{P}_{\sigma_1,\sigma_2}$ be the path from σ_1 to σ_2 constructed as follows:

- Update the vertices $v_i \in V$ in order specified by $\prec$.
- If $\lambda > 0$ then update v_i by recolouring v_i with colour $\sigma_2(v_i)$.
- If $\lambda = 0$ then update v_i in the following way:
  1. Let W be the set of neighbours of v_i in R that have colour $\sigma_2(v_i)$. The vertices in W do not have their their final colour in σ_2 yet. If W is non-empty, recolour the vertices $w_i \in W$ (in order specified by $\prec$) by recolouring w_i with the smallest available colour in Q, such that the colouring of R stays in $\Omega_R(\mathcal{B})$. There is always such a colour since we assume $q \geq \Delta + 2$ if $\lambda = 0$.
  2. Recolour v_i with colour $\sigma_2(v_i)$.

Assign all of the flow from σ_1 to σ_2 to path $\gamma_{\sigma_1,\sigma_2}$. That is, $\phi(\gamma_{\sigma_1,\sigma_2}) = P_d(\sigma_1, \sigma_2)\pi_{\mathcal{B}}(\sigma_1)$, and $\phi(\gamma) = 0$ for all paths $\gamma \in \mathcal{P}_{\sigma_1,\sigma_2}$ unless $\gamma = \gamma_{\sigma_1,\sigma_2}$. Let θ_1 and θ_2, where $(\theta_1, \theta_2) \in E_0$, be two colourings that disagree on a vertex w. Then the congestion of edge (θ_1, θ_2)

is

$$\begin{aligned}
A_{\theta_1,\theta_2}(\phi) &= \frac{1}{\pi_{\mathcal{B}}(\theta_1)P_0(\theta_1,\theta_2)} \sum_{\gamma\in\mathcal{P}:(\theta_1,\theta_2)\in\gamma} |\gamma|\phi(\gamma) \\
&= \frac{1}{\pi_{\mathcal{B}}(\theta_1)P_0(\theta_1,\theta_2)} \sum_{\substack{(\sigma_1,\sigma_2)\in E_d:\\ (\theta_1,\theta_2)\in\gamma_{\sigma_1,\sigma_2}}} |\gamma_{\sigma_1,\sigma_2}| P_d(\sigma_1,\sigma_2)\pi_{\mathcal{B}}(\sigma_1) \\
&= \sum_{\substack{(\sigma_1,\sigma_2)\in E_d:\\ (\theta_1,\theta_2)\in\gamma_{\sigma_1,\sigma_2}}} |\gamma_{\sigma_1,\sigma_2}| \cdot \frac{P_d(\sigma_1,\sigma_2)}{P_0(\theta_1,\theta_2)} \cdot \frac{\pi_{\mathcal{B}}(\sigma_1)}{\pi_{\mathcal{B}}(\theta_1)} \\
&\leq k_1 \cdot k_2 \cdot \sum_{\substack{(\sigma_1,\sigma_2)\in E_d:\\ (\theta_1,\theta_2)\in\gamma_{\sigma_1,\sigma_2}}} \frac{P_d(\sigma_1,\sigma_2)}{P_0(\theta_1,\theta_2)} \\
&\leq k_1 \cdot k_2 \cdot k_3 \leq O(1),
\end{aligned}$$

where k_1, k_2 and k_3 are constants, specified next.

The path length $|\gamma_{\sigma_1,\sigma_2}|$ is upper-bounded by a constant k_1 since σ_1 and σ_2 differ only on vertices inside a ball of fixed radius d. The path $\gamma_{\sigma_1,\sigma_2}$ is constructed such that only vertices on which σ_1 and σ_2 differ are recoloured, and they are recoloured at most a constant number of times.

The fraction $\pi_{\mathcal{B}}(\sigma_1)/\pi_{\mathcal{B}}(\theta_1)$ is upper-bounded by a constant k_2 since σ_1 and θ_1 differ only on vertices inside a ball of fixed radius d. Let m_{σ_1} be the number of monochromatic edges in $E(R)$ under σ_1 and $\mathcal{B}$, and let m_{θ_1} be the number of monochromatic edges in $E(R)$ under θ_1 and $\mathcal{B}$. Hence m_{σ_1} and m_{θ_1}differ by at most some constant. Thus, $\pi_{\mathcal{B}}(\sigma_1)/\pi_{\mathcal{B}}(\theta_1) = \lambda^{m_{\sigma_1}}/\lambda^{m_{\theta_1}}$ is upper-bounded by a constant k_2.

Lastly, to see that the sum is bounded by a constant k_3, note that there are only a constant number of pairs (σ_1,σ_2) in the summation. This holds since σ_1 and σ_2 agree with θ_1 on all vertices in R except in a constant-sized ball around vertex w on which θ_1 and θ_2 differ. We have

$$P_d(\sigma_1,\sigma_2) \leq \frac{m}{|R^d|} \in O(\frac{1}{|R^d|}),$$

where m is the number of vertices for which the heat-bath update in $\mathcal{M}_d$ will update all vertices on which σ_1 and σ_2 differ. Note that m is bounded by a constant since σ_1 and σ_2 differ only on vertices inside a ball of fixed radius d. Furthermore, we have

$$P_0(\theta_1,\theta_2) \geq \frac{1}{|R^d|} \cdot \frac{\lambda^\Delta}{q} \in \Omega(\frac{1}{|R^d|})$$

since λ^Δ/q is the smallest probability of making a transition in $\mathcal{M}_0$ from colouring θ_1 to θ_2 once vertex w on which θ_1 and θ_2 differ has been chosen for a heat-bath update.

Thus,

$$\frac{P_d(\sigma_1, \sigma_2)}{P_0(\theta_1, \theta_2)} \in O(1)$$

and we have that the sum is bounded by a constant k_3.

Now, $A_{\theta_1,\theta_2}(\phi) \in O(1)$ for all $(\theta_1, \theta_2) \in E_0$ and it follows that the congestion $A(\phi) \in O(1)$. □

Finally we have the machinery for proving Theorem 2.15. We repeat the theorem and give the proof.

Theorem (2.15). *Consider the anti-ferromagnetic Potts model on a uniformly sub-exponential graph $\mathcal{G}$ with maximum degree Δ, and parameters q and λ. If the spin system has a K-coupling cover, and $\lambda > 0$ or $q \geq \Delta + 2$, then the Glauber dynamics is rapidly mixing for any region R and any q_0-colouring $\mathcal{B}$ of ∂R. More precisely, the mixing time $\tau(\delta) \leq O(n(n + \log \frac{1}{\delta}))$, where n is the number of vertices in R.*

Proof. From Lemma 4.5 it follows that there is an integer $d \geq 0$ such that $\mathcal{M}_d$ has mixing time $\tau_{\mathcal{M}_d}(\delta) \in O(n \log \frac{n}{\delta})$. Then Lemma 4.7 together with Lemma 4.8 gives the desired mixing time $\tau_{\mathcal{M}_0}(\delta) \in O(n(n + \log \frac{1}{\delta}))$ of the Glauber dynamics $\mathcal{M}_0$. □

We make a final remark that the mixing bound $O(n(n + \log \frac{1}{\delta}))$ of the Glauber dynamics in Theorem 2.15 is not optimal. From Lemma 4.5 we have that the mixing time of $\mathcal{M}_d$ is $O(n \log \frac{n}{\delta})$ for some $d > 0$. The difference in mixing time is due to the general Markov chain comparison technique we use. It might be possible to reduce the difference by considering a more graph-colouring specific comparison between $\mathcal{M}_d$ and the Glauber dynamics.

It is known that for some graphs (in particular the square lattice [GJMP06]) strong spatial mixing implies $O(n \log \frac{n}{\delta})$ mixing of the Glauber dynamics. As explained in [DSVW04], this can be proved using techniques from functional analysis. Functional analysis has been used to prove that strong spatial mixing implies optimal mixing of the Glauber dynamics in the general case [Ces01, Mar99, MO94a, MO94b, SZ92], however, a purely combinatorial proof of this implication has only been given for *monotone spin systems* [DSVW04].

Chapter 5

Boundary pairs

When Goldberg, Martin and Paterson showed strong spatial mixing in [GMP05], they introduced the concept of a *boundary pair*. Boundary pairs are data structures holding information about a region R, colourings of the boundary of R, etc. There are two types of boundary pairs: *vertex-boundary pairs* and *edge-boundary pairs*. A vertex-boundary pair contains information about the colouring of the vertex boundary of a region. The idea is to derive certain properties of the vertex-boundary pairs which can easily be translated into properties of the spin system, such as whether there is strong spatial mixing or not. When Goldberg, Martin and Paterson [GMP05] derived these properties, it turned out to be convenient to work with edge-boundary pairs. An edge-boundary pair contains colourings of the edge boundary $\mathcal{E}R$ rather than the vertex boundary ∂R. This chapter will define a vertex- and edge-boundary pair and describe how they are related. The definition of an edge-boundary pair is slightly modified from the definition in [GMP05] in order to suit the contents of this thesis. In this chapter we will also see how certain properties of the edge-boundary pairs imply that the system has strong spatial mixing and a K-coupling cover. We use the calligraphic font ($\mathcal{X}$, $\mathcal{B}$, etc.) when referring to a vertex-boundary pair, and the normal font (X, B, etc.) when referring to an edge-boundary pair.

The chapter is organised as follows. In Section 5.1 we define the notion of a vertex-boundary pair. In Section 5.2 we define the notion of an edge-boundary pair and explain how it is related to a vertex-boundary pair. In Section 5.3 we present some lemmas stating that if certain conditions of the edge-boundary pairs are satisfied then the system has strong spatial mixing and a K-coupling cover.

5.1 Vertex-boundary pairs

For a fixed infinite graph $\mathcal{G}$ and a fixed number of colours q, a vertex-boundary pair is defined as follows.

Definition 5.1 (Vertex-boundary pair). *A* vertex-boundary pair $\mathcal{X}$ *consists of*

- *a region* $R_{\mathcal{X}}$,
- *a distinguished boundary vertex* $w_{\mathcal{X}} \in \partial R_{\mathcal{X}}$, *and*
- *a pair* $(\mathcal{B}_{\mathcal{X}}, \mathcal{B}'_{\mathcal{X}})$ *of* q_0*-colourings of* $\partial R_{\mathcal{X}}$ *that are identical on all vertices except on* $w_{\mathcal{X}}$, *where they differ. The colour of* $w_{\mathcal{X}}$ *is in* Q *for both* $\mathcal{B}_{\mathcal{X}}$ *and* $\mathcal{B}'_{\mathcal{X}}$.

Note that the colour of the distinguished vertex $w_{\mathcal{X}}$ has to be in the set Q. That is, $w_{\mathcal{X}}$ cannot be a free-boundary vertex with colour 0. Now, Definition 2.10 of strong spatial mixing can be rephrased using the definition of a vertex-boundary pair: The anti-ferromagnetic Potts model on the infinite graph $\mathcal{G}$ has strong spatial mixing for parameters λ and q if there are constants $\alpha > 0$ and $0 < \varepsilon < 1$ such that, for every vertex-boundary pair $\mathcal{X}$ and every $\Lambda \subseteq R_{\mathcal{X}}$,

$$\mathrm{d_{TV}}(\pi_{\mathcal{B}_{\mathcal{X}},\Lambda}, \pi_{\mathcal{B}'_{\mathcal{X}},\Lambda}) \leq \alpha|\Lambda|(1-\varepsilon)^{d(w_{\mathcal{X}},\Lambda)}.$$

We are interested in the effect that the difference in the colour of $w_{\mathcal{X}}$ has on the vertices in $R_{\mathcal{X}}$. Recall the definition of a coupling in Section 2.2, and let $\Psi_{\mathcal{X}}$ be a coupling of $\pi_{\mathcal{B}_{\mathcal{X}}}$ and $\pi_{\mathcal{B}'_{\mathcal{X}}}$. For a vertex $v \in R_{\mathcal{X}}$ we define the indicator random variable $1_{\Psi_{\mathcal{X}},v}$ for the event that the colour of v differs in a pair of colourings drawn from $\Psi_{\mathcal{X}}$. Hence, the quantity $\sum_{v \in R_{\mathcal{X}}} \mathbb{E}[1_{\Psi_{\mathcal{X}},v}]$ is the expected number of vertices in $R_{\mathcal{X}}$ on which the colour differs in a pair of colourings drawn from $\Psi_{\mathcal{X}}$. If $\mathbb{E}[1_{\Psi_{\mathcal{X}},v}]$ is small enough for all vertex-boundary pairs $\mathcal{X}$ and vertices $v \in R_{\mathcal{X}}$, then we can use that conclusion to infer strong spatial mixing. This will be proved in Lemma 5.4. Furthermore, if $\sum_{v \in R_{\mathcal{X}}} \mathbb{E}[1_{\Psi_{\mathcal{X}},v}]$ is upper-bounded by a constant K then it follows that the system has a K-coupling cover (Definition 2.11). One way to show that the sum it small is to show that $\mathbb{E}[1_{\Psi_{\mathcal{X}},v}]$ decreases rapidly as the distance between $w_{\mathcal{X}}$ and v increases. We want to construct a coupling $\Psi_{\mathcal{X}}$ such that $\mathbb{E}[1_{\Psi_{\mathcal{X}},v}]$ decreases exponentially with the distance between $w_{\mathcal{X}}$ and v, and then show that $\sum_{v \in R_{\mathcal{X}}} \mathbb{E}[1_{\Psi_{\mathcal{X}},v}]$ is upper-bounded by a constant.

5.2 Edge-boundary pairs

When constructing the coupling $\Psi_{\mathcal{X}}$ of $\pi_{\mathcal{B}_{\mathcal{X}}}$ and $\pi_{\mathcal{B}'_{\mathcal{X}}}$ for a vertex-boundary pair, it is convenient to work with edge boundaries. Similarly to the definition of a vertex colouring we define a q-, q_0- and 0-colouring of a set $E \subseteq E_{\mathcal{G}}$ of edges to be a function from E to Q, Q_0 and $\{0\}$, respectively. If B is a q_0-colouring of E, and E' is a subset of E then $B(E')$ is the colouring of E' induced by B. Furthermore, $B(e)$ is the colour of edge $e \in E$ under B.

Let σ be a colouring of a region R, let $\mathcal{B}$ be a colouring of ∂R, let B be a colouring of the boundary edges $\mathcal{E}R$ and let $E \subseteq E(R)$ be a set of edges. Analogous to the definition of $\text{mon}_{\sigma,\mathcal{B}}(E)$ in Section 2.1.2, where $\mathcal{B}$ is a q_0-colouring of ∂R, $\text{mon}_{\sigma,B}(E)$ is the number of monochromatic edges in E under σ and B. An edge $e \in \mathcal{E}R$ is defined to be monochromatic if $B(e) = \sigma(v)$, where vertex $v \in R$ is incident to e. The distribution π_B on vertex colourings of R is defined in the same way as $\pi_{\mathcal{B}}$ in Section 2.1.2, only with $\mathcal{B}$ replaced by B. The two colourings $\mathcal{B}$ and B are *colour-equivalent* if, and only if, for every edge $e = (u, v) \in \mathcal{E}R$, where vertex $u \in \partial R$, the colour $B(e) = \mathcal{B}(u)$. If $\mathcal{B}$ and B are colour-equivalent it follows from the definitions of monochromatic edges that the distributions $\pi_{\mathcal{B}}$ and π_B are identical.

Consider a vertex-boundary pair $\mathcal{X}$. There is a pair (B, B') of colourings of $\mathcal{E}R_{\mathcal{X}}$ such that $\mathcal{B}_{\mathcal{X}}$ and B, and $\mathcal{B}'_{\mathcal{X}}$ and B' are colour-equivalent. We define a coupling $\Psi_{\mathcal{X}}$ of $\pi_{\mathcal{B}_{\mathcal{X}}}$ and $\pi_{\mathcal{B}'_{\mathcal{X}}}$ (or π_B and $\pi_{B'}$, equivalently) by composing couplings $\Psi_1, \ldots, \Psi_k$ of distributions derived from edge-boundary colourings that differ on only one edge. Here $k = |E| \geq 1$, where $E = \{e_1, \ldots, e_k\} \subseteq \mathcal{E}R_{\mathcal{X}}$ is the set of boundary edges incident to $w_{\mathcal{X}} \in \partial R_{\mathcal{X}}$. We do not consider any particular ordering of the edges in E. For $i = 0, \ldots, k$, let B_i be the colouring of $\mathcal{E}R_{\mathcal{X}}$ such that

- $B_i(e) = B(e)$ (and hence $B_i(e) = B'(e)$) if $e \in \mathcal{E}R_{\mathcal{X}} \backslash E$,
- $B_i(e_j) = \mathcal{B}'_{\mathcal{X}}(w_{\mathcal{X}})$ if $j \leq i$, and
- $B_i(e_j) = \mathcal{B}_{\mathcal{X}}(w_{\mathcal{X}})$ if $j > i$,

where $j = 1, \ldots, k$ and $e_j \in E$. In other words, the colouring B_0 is identical to B, and the colouring B_k is identical to B'. Two colourings B_{i-1} and B_i differ only on edge $e_i \in E$; $B_{i-1}(e_i) = \mathcal{B}_{\mathcal{X}}(w_{\mathcal{X}})$ and $B_i(e_i) = \mathcal{B}'_{\mathcal{X}}(w_{\mathcal{X}})$. An example with a 3-vertex region $R_{\mathcal{X}}$ of the triangular lattice is illustrated in Figure 5.1. Here the vertex $w_{\mathcal{X}}$ on the boundary is shaded.

For $i = 1, \ldots, k$, let Ψ_i denote a coupling of $\pi_{B_{i-1}}$ and π_{B_i}. A coupling $\Psi_{\mathcal{X}}$ of $\pi_{\mathcal{B}_{\mathcal{X}}}$ and $\pi_{\mathcal{B}'_{\mathcal{X}}}$ (or π_{B_0} and π_{B_k}, equivalently) is constructed by composing the couplings $\Psi_1, \ldots, \Psi_k$ in the natural way. To choose a pair (σ_0, σ_k) from $\Psi_{\mathcal{X}}$ we first choose the pair (σ_0, σ_1) from Ψ_1. Say $\sigma_0 = x_0$ and $\sigma_1 = x_1$. Then choose the pair (σ_1, σ_2) from the conditional distribution Ψ_2, conditioned on $\sigma_1 = x_1$. Say that $\sigma_2 = x_2$. Then choose the pair (σ_2, σ_3) from the conditional distribution Ψ_3, conditioned on $\sigma_2 = x_2$, and so on. Hence, σ_0 is drawn from π_{B_0} and σ_k is drawn from π_{B_k}.

For a vertex $v \in R_{\mathcal{X}}$ and a coupling Ψ_i, define the indicator random variable $1_{\Psi_i, v}$ for the event that the colour of v differs in a pair of colourings drawn from Ψ_i. From the construction of the coupling $\Psi_{\mathcal{X}}$ it follows that if the colour of v differs in a pair (σ_0, σ_k) drawn from $\Psi_{\mathcal{X}}$ then it must differ in at least one of the pairs (σ_{i-1}, σ_i)

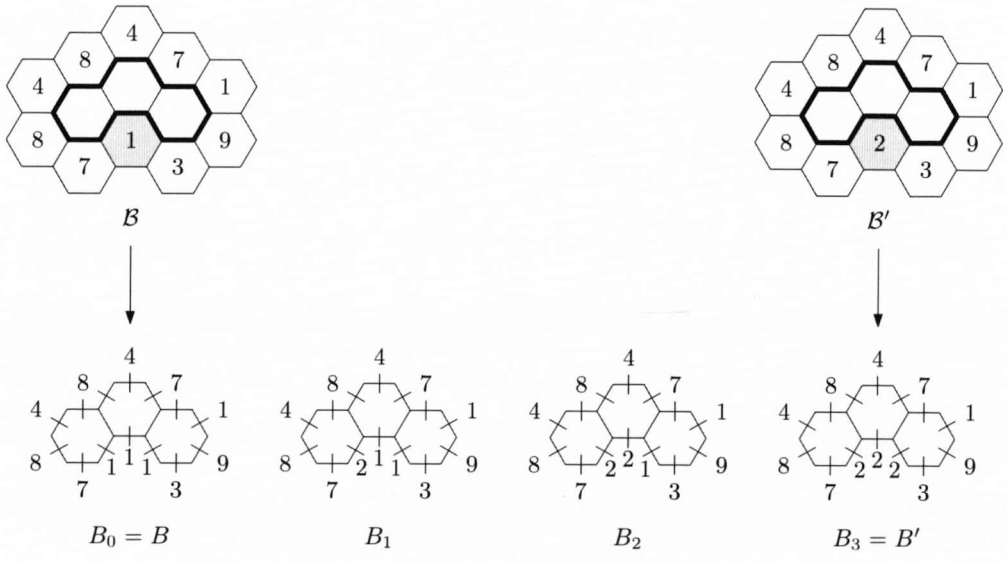


Figure 5.1. A 3-vertex region R with two colourings $\mathcal{B}$ and $\mathcal{B}'$ of ∂R that differ only on the shaded vertex. The edge colouring B_0 of $\mathcal{E}R$ is colour-equivalent to $\mathcal{B}$, and the edge colouring B_3 is colour-equivalent to $\mathcal{B}'$. A pair (B_{i-1}, B_i) of edge colourings, $i = 1, \ldots, 3$, differ only on one boundary edge, and Ψ_i denotes a coupling of $\pi_{B_{i-1}}$ and π_{B_i}.

drawn from Ψ_i, where $i = 1, \ldots, k$. Thus,

$$\mathbb{E}[1_{\Psi_{\mathcal{X}}, v}] \leq \sum_{i=1}^{k} \mathbb{E}[1_{\Psi_i, v}].$$

If, for $i = 1, \ldots, k$, the quantity $\mathbb{E}[1_{\Psi_i, v}]$ decreases exponentially with the distance between edge e_i (on which B_{i-1} and B_i differ) and vertex v, then the quantity $\mathbb{E}[1_{\Psi_{\mathcal{X}}, v}]$ decreases exponentially with the distance between $w_{\mathcal{X}}$ and v. This is the property we would like to establish. Thus, the difference in the colour of a single vertex on the boundary is broken into differences in the edges that bound it. It is mainly in this setting of edge-discrepancies we will operate.

Before introducing the notion of an edge-boundary pair, we define *adjacent edges* for planar graphs. Let $\mathcal{G}$ be a planar graph with a fixed layout (such that no edges intersect). This layout of $\mathcal{G}$ specifies a unique clockwise ordering of the edges that are incident to a vertex $u \in V_{\mathcal{G}}$. Any two edges $e = (u, v)$ and $e' = (u, v')$ in $E_{\mathcal{G}}$ that are incident to the same vertex u are adjacent if e follows directly after e' or if e' follows directly after e when traversing the edges that are incident to u clockwise around u. The degree of a vertex v is denoted $\deg(v)$.

Definition 5.2 (Edge-boundary pair). *An* edge-boundary pair X *consists of*

- *a region* R_X,
- *a distinguished boundary edge* $e_X = (w_X, v_X) \in \mathcal{E}R_X$ *with* $w_X \in \partial R_X$, $v_X \in R_X$, *and*

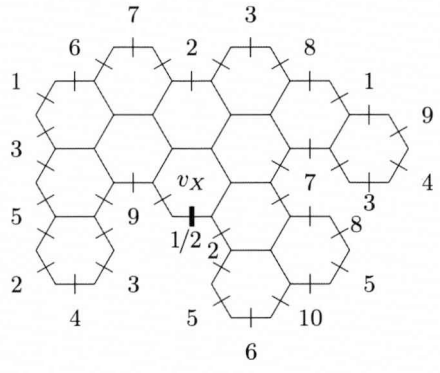


Figure 5.2. An example of how an edge-boundary pair X is illustrated. The edge boundary is labelled with its colouring. The edge e_X is marked with a thick line segment and labelled "$B_X(e_X)/B'_X(e_X)$".

- *a pair (B_X, B'_X) of q_0-colourings of $\mathcal{E}R_X$ that are identical on all edges except on e_X, where they differ. The colour of e_X is in Q for both B_X and B'_X.*

Furthermore, if the underlying graph $\mathcal{G}$ is planar, we require

- *any two adjacent boundary edges that share a vertex in ∂R_X have the same colour in at least one of the two colourings B_X and B'_X (and so in both of B_X and B'_X except when edge e_X is involved), and*

- *if $\deg(w_X) \geq 2$ then there is a vertex adjacent to w_X which is not in R_X.*

The definition of an edge-boundary pair X is the analogue of the definition of a vertex-boundary pair $\mathcal{X}$, with the discrepancy on edge e_X instead of vertex $w_{\mathcal{X}}$. For an edge-boundary pair X, let n_X denote the number of neighbours of v_X that are in R_X.

When illustrating edge-boundary pairs X in the figures, we mark the edges in the edge boundary with short line segments. The edge e_X is marked with a short and thick line segment, and we often label vertex v_X. The edges in the edge boundary are labelled with their colours in B_X and B'_X, and edge e_X is labelled "$B_X(e_X)/B'_X(e_X)$". Figure 5.2(a) illustrates an edge-boundary pair X on the triangular lattice, where the colour of e_X is 1 in B_X and 2 in B'_X.

In the definition of an edge-boundary pair there are two conditions that have to be satisfied if the graph $\mathcal{G}$ is planar. The first condition imposes a natural restriction on possible colourings B_X and $B_{X'}$ of $\mathcal{E}R_X$, given a region R_X. Edge colourings are derived from vertex colourings, and hence it makes sense to consider edge colourings that are colour-equivalent to colourings of the vertex boundary. For instance, consider edge-boundary pairs X where R_X is a region of the triangular lattice (which is planar), and the colour of edge e_X is 1 and 2 in B_X and B'_X, respectively. Figure 5.3(a) illustrates a colouring of the edge boundary of a valid edge-boundary pair X. The whole region is not shown. Figure 5.3(b) and (c) on the other hand, illustrate colourings of

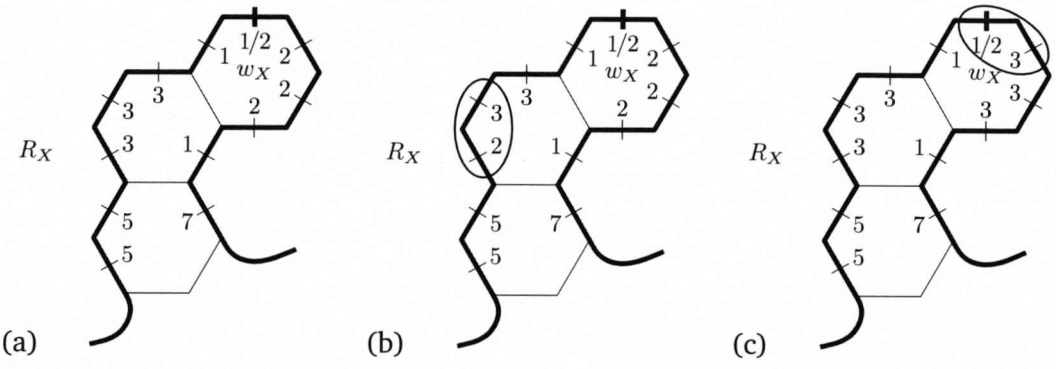


Figure 5.3. A part of the boundary of a region R_X of an edge-boundary pair X. Note that edge e_X has colour 1 in B_X and colour 2 in B'_X. (a) A valid colouring of the boundary edges. (b) An invalid colouring since the two adjacent edges that are encircled have received different colours. (c) An invalid colour since edge e_X does not have colour 3 in either B_X or B'_X.

the edge boundary that are not allowed for any edge-boundary pair X. The pairs of adjacent edges violating the constraint in the definition are encircled in the figures. The constraint clearly reduces the number of edge-boundary pairs X for a fixed region R_X and a fixed number of colours q. This is of great importance in the computer-assisted proofs, in which we exhaustively search sets of edge-boundary pairs defined for fixed regions R_X.

The last constraint in the definition of an edge-boundary pair makes sure that vertex w_X is not a "hole" in the region R_X. As we will see in subsequent chapters, this restriction lets us avoid treating several technicalities in the computer-assisted proofs of exponential decay.

5.3 Strong spatial mixing and K-coupling cover

One way of showing strong spatial mixing for a spin system is to show exponential decay for every edge-boundary pair X. Exponential decay for every edge-boundary pair will imply exponential decay for every vertex-boundary pair $\mathcal{X}$ (Lemma 5.3), which in turn implies strong spatial mixing (Lemma 5.4) and that there is a K-coupling cover (Lemma 5.5). The three lemmas are stated next.

Lemma 5.3. *If, for all edge-boundary pairs X, there is a coupling Ψ of π_{B_X} and $\pi_{B'_X}$ such that*

$$\mathbb{E}[1_{\Psi,v}] \leq \alpha(1-\varepsilon)^{d(e_X,v)},$$

for every vertex $v \in R_X$, where $\alpha > 0$ and $0 < \varepsilon < 1$ are constants, then, for every vertex-boundary pair $\mathcal{X}$, there is a coupling $\Psi_{\mathcal{X}}$ of $\pi_{\mathcal{B}_{\mathcal{X}}}$ and $\pi_{\mathcal{B}'_{\mathcal{X}}}$ such that

$$\mathbb{E}[1_{\Psi_{\mathcal{X}},v}] \leq \alpha'(1-\varepsilon)^{d(w_{\mathcal{X}},v)},$$

for all vertices $v \in R_{\mathcal{X}}$, where $\alpha' > 0$ is a constant.

Proof. Let $\mathcal{X}$ be any vertex-boundary pair. First suppose that $w_{\mathcal{X}}$ has a neighbour $y \notin R_{\mathcal{X}}$. Let $k = |E|$, where $E = \{e_1, \ldots, e_k\} \subseteq \mathcal{E}R_{\mathcal{X}}$ is the set of boundary edges incident to $w_{\mathcal{X}}$. If $\mathcal{G}$ is planar, label the edges in E clockwise around $w_{\mathcal{X}}$ so that edge $(w_{\mathcal{X}}, y)$ appears between edge e_k and e_1 when traversing edges around $w_{\mathcal{X}}$ in clockwise direction. This guarantees that e_i and e_j are adjacent only if i and j differ by 1. If $\mathcal{G}$ is not planar, label the edges in E in any well-defined order.

Let B be the edge-boundary colouring of $\mathcal{E}R_{\mathcal{X}}$ that is colour-equivalent to $\mathcal{B}_{\mathcal{X}}$. For $i = 1, \ldots, k$, let X_i be the edge-boundary pair consisting of region $R_{X_i} = R_{\mathcal{X}}$, the distinguished edge $e_{X_i} = (w_{X_1}, v_{X_i}) = e_i$, and boundary colourings B_{X_i} and B'_{X_i} that are both identical to B on all edges in $\mathcal{E}R_{X_i} \backslash E$. The colours of the edges in E are assigned as follows.

- $B_{X_i}(e_j) = B'_{X_i}(e_j) = \mathcal{B}'_{\mathcal{X}}(w_{\mathcal{X}})$ for $j = 1, \ldots, i-1$,
- $B_{X_i}(e_j) = \mathcal{B}_{\mathcal{X}}(w_{\mathcal{X}})$ and $B'_{X_i}(e_j) = \mathcal{B}'_{\mathcal{X}}(w_{\mathcal{X}})$ for $j = i$, and
- $B_{X_i}(e_j) = B'_{X_i}(e_j) = \mathcal{B}_{\mathcal{X}}(w_{\mathcal{X}})$ for $j = i+1, \ldots, k$.

Figure 5.1 on page 52 can be used as an illustration of how the sequence of edge-boundary colourings is derived. Let Ψ_i be a coupling of $\pi_{B_{X_i}}$ and $\pi_{B'_{X_i}}$ such that $\mathbb{E}[1_{\Psi_i, v}] \leq \alpha(1-\varepsilon)^{d(e_{X_i}, v)}$ for all vertices $v \in R_{X_i}$. Let $\Psi_{\mathcal{X}}$ be the coupling of $\pi_{\mathcal{B}_{\mathcal{X}}}$ and $\pi_{\mathcal{B}'_{\mathcal{X}}}$ defined by composing the couplings Ψ_i for $i = 1, \ldots, k$. See the previous section for details on this composition. It follows that

$$\mathbb{E}[1_{\Psi_{\mathcal{X}}, v}] \leq \sum_{i=1}^{k} \mathbb{E}[1_{\Psi_i, v}] \leq (\Delta - 1)\alpha(1-\varepsilon)^{d(w_{\mathcal{X}}, v)},$$

since $k \leq \Delta - 1$.

Now we must deal with the case in which all neighbours of $w_{\mathcal{X}}$ are in $R_{\mathcal{X}}$. Breaking the discrepancy at vertex $w_{\mathcal{X}}$ into edge-boundary pairs X_i as above is no longer straightforward if $\mathcal{G}$ is planar and $\deg(w_{\mathcal{X}}) \geq 2$. From the definition of an edge-boundary pair X there must be a neighbour of w_X that is not in R_X.

Let $u \in R_{\mathcal{X}}$ be any neighbour of $w_{\mathcal{X}}$. Let $R_{\mathcal{X},u} = R_{\mathcal{X}} \backslash \{u\}$ be the region $R_{\mathcal{X}}$ after removing vertex u. Let $\mathcal{B}_{\mathcal{X},c}$ be the colouring of the vertex-boundary $\partial R_{\mathcal{X},u}$ such that $\mathcal{B}_{\mathcal{X},c}(v) = \mathcal{B}_{\mathcal{X}}(v)$ on all vertices $v \in \partial R_{\mathcal{X}} \cap \partial R_{\mathcal{X},u}$, and $\mathcal{B}_{\mathcal{X},c}(u) = c$ (if $u \in \partial R_{\mathcal{X},u}$), where $c \in Q$. Similarly, let $\mathcal{B}'_{\mathcal{X},c'}$ be the colouring of the vertex-boundary $\partial R_{\mathcal{X},u}$ such that $\mathcal{B}'_{\mathcal{X},c'}(v) = \mathcal{B}'_{\mathcal{X}}(v)$ on all vertices $v \in \partial R_{\mathcal{X}} \cap \partial R_{\mathcal{X},u}$, and $\mathcal{B}'_{\mathcal{X},c'}(u) = c'$, where $c' \in Q$. Given two colours c and c', note that the colourings $\mathcal{B}_{\mathcal{X},c}$ and $\mathcal{B}'_{\mathcal{X},c'}$ can differ on up to two vertices, namely vertex $w_{\mathcal{X}}$ and u. We break the difference in the (up to) two vertices $w_{\mathcal{X}}$ and u on the boundary $\partial R_{\mathcal{X},u}$ into differences in the edges that bound them.

For two (not necessarily different) fixed colours c and c', let B be the edge-colouring of $\mathcal{E}R_{\mathcal{X},u}$ that is colour-equivalent to $\mathcal{B}_{\mathcal{X},c}$, and let B' be the edge-colouring of $\mathcal{E}R_{\mathcal{X},u}$ that is colour-equivalent to $\mathcal{B}'_{\mathcal{X},c'}$. Let $k = |E|$, where $E = \{e_1, \ldots, e_k\} \subseteq \mathcal{E}R_{\mathcal{X},u}$ be the set of boundary edges that receive different colours in B and B'. Note that edges in E can only be incident to $w_{\mathcal{X}}$ and u. Thus, k is at most $2\Delta - 2$. If $\mathcal{G}$ is planar, label the edges in E clockwise around $w_{\mathcal{X}}$ and u so that e_k and e_1 are not adjacent. Such a labelling is always possible (since $w_{\mathcal{X}}$ and u are neighbours) and guarantees that if e_i and e_j share a vertex in $\partial R_{\mathcal{X},u}$ then e_i and e_j are only adjacent if i and j differ by 1. If $\mathcal{G}$ is not planar, label the edges in E in any well-defined order. Similarly to before, for $i = 1, \ldots, k$, let X_i be the edge-boundary pair consisting of region $R_{X_i} = R_{\mathcal{X},u}$, the distinguished edge $e_{X_i} = e_i$, and boundary colourings B_{X_i} and B'_{X_i} defined similarly to before, as a sequence of colourings going from B to B', differing only on the distinguished edge e_i. Let Ψ_i be a coupling of $\pi_{B_{X_i}}$ and $\pi_{B'_{X_i}}$ such that $\mathbb{E}[1_{\Psi_i,v}] \leq \alpha(1-\varepsilon)^{d(e_{X_i},v)}$ for all vertices $v \in R_{X_i}$. We now construct a coupling $\Psi_{\mathcal{X}}$ of $\pi_{\mathcal{B}_{\mathcal{X}}}$ and $\pi_{\mathcal{B}'_{\mathcal{X}}}$ in the following way.

Let $\Psi'_{\mathcal{X}}$ be *any* coupling of $\pi_{\mathcal{B}_{\mathcal{X}}}$ and $\pi_{\mathcal{B}'_{\mathcal{X}}}$. Let (C, C') be the random variable corresponding to the pair of colourings drawn from $\Psi_{\mathcal{X}}$ (yet to be constructed). We will choose the colour of u in C and C' according to $\Psi'_{\mathcal{X}}$. Let c and c' be the colour of u drawn from $\Psi'_{\mathcal{X}}$. To complete the construction of $\Psi_{\mathcal{X}}$ we colour the remaining vertices in $R_{\mathcal{X}}$ by choosing two colourings from $\Psi_{\mathcal{X},c,c'}$, where $\Psi_{\mathcal{X},c,c'}$ is a coupling of $\pi_{\mathcal{B}_{\mathcal{X},c}}$ and $\pi_{\mathcal{B}'_{\mathcal{X},c'}}$. The coupling $\Psi_{\mathcal{X},c,c'}$ is constructed by composing the k couplings Ψ_{X_i}. For $v \neq u$ it follows that

$$\mathbb{E}[1_{\Psi_{\mathcal{X}},v}] \leq \sum_{i=1}^{k} \mathbb{E}[1_{\Psi_i,v}] \leq (2\Delta-2)\alpha(1-\varepsilon)^{d(w_{\mathcal{X}},v)-1} = \frac{(2\Delta-2)\alpha}{(1-\varepsilon)}(1-\varepsilon)^{d(w_{\mathcal{X}},v)},$$

since $k \leq 2\Delta - 2$. The -1 in the exponent comes from the fact that the distance from the discrepancy edge to a vertex v may be one less than $d(w_{\mathcal{X}}, v)$. For $v = u$ we have $d(w_{\mathcal{X}}, v) = 1$ so by choosing $\alpha' \geq 1/(1-\varepsilon)$ we ensure that $\mathbb{E}[1_{\Psi_{\mathcal{X}},v}] \leq \alpha'(1-\varepsilon)$. □

Lemma 5.4 (Strong spatial mixing). *If, for all vertex-boundary pairs $\mathcal{X}$, there is a coupling $\Psi_{\mathcal{X}}$ of $\pi_{\mathcal{B}_{\mathcal{X}}}$ and $\pi_{\mathcal{B}'_{\mathcal{X}}}$ such that*

$$\mathbb{E}[1_{\Psi_{\mathcal{X}},v}] \leq \alpha(1-\varepsilon)^{d(w_{\mathcal{X}},v)},$$

for every vertex $v \in R_{\mathcal{X}}$, where $\alpha > 0$ and $0 < \varepsilon < 1$ are constants, then the spin system has strong spatial mixing.

Proof. Consider the vertex-boundary pair $\mathcal{X}$ such that, from Definition 2.10 of strong spatial mixing, we have $R_{\mathcal{X}} = R$, $\mathcal{B}_{\mathcal{X}} = \mathcal{B}$, $\mathcal{B}'_{\mathcal{X}} = \mathcal{B}'$ and $w_{\mathcal{X}} = w$. The total variation distance between $\pi_{\mathcal{B},\Lambda}$ and $\pi_{\mathcal{B}',\Lambda}$ is upper-bounded by the probability that Λ dif-

fers under any coupling Ψ of $\pi_{\mathcal{B},\Lambda}$ and $\pi_{\mathcal{B}',\Lambda}$. This probability is upper-bounded by $\sum_{v\in\Lambda}\mathbb{E}[1_\Psi, v]$. With $\Psi = \Psi_\mathcal{X}$ we have

$$d_{\text{tv}}(\pi_{\mathcal{B},\Lambda}, \pi_{\mathcal{B}',\Lambda}) \leq \sum_{v\in\Lambda}\mathbb{E}[1_{\Psi_\mathcal{X}}, v] \leq |\Lambda|\alpha(1-\varepsilon)^{d(w_\mathcal{X},\Lambda)},$$

which is the desired upper bound. □

Lemma 5.5 (K-coupling cover). *Suppose $\mathcal{G}$ is an infinite graph that is uniformly sub-exponential. If, for all vertex-boundary pairs $\mathcal{X}$, there is a coupling $\Psi_\mathcal{X}$ of $\pi_{\mathcal{B}_\mathcal{X}}$ and $\pi_{\mathcal{B}'_\mathcal{X}}$ such that*

$$\mathbb{E}[1_{\Psi_\mathcal{X},v}] \leq \alpha(1-\varepsilon)^{d(w_\mathcal{X},v)},$$

for every vertex $v \in R_\mathcal{X}$, where $\alpha > 0$ and $0 < \varepsilon < 1$ are constants, then the spin system has a K-coupling cover, where K is a constant that depends on the spin system, α and ε.

Proof. Let $\mathcal{X}$ be a vertex-boundary pair. Let $n_d(w_\mathcal{X})$ be the number of vertices in $V_\mathcal{G}$ that are at distance d from $w_\mathcal{X}$. We will show that

$$\sum_{v\in R_\mathcal{X}}\mathbb{E}[1_{\Psi_\mathcal{X},v}] \leq K,$$

where $\Psi_\mathcal{X}$ is a coupling of $\pi_{\mathcal{B}_\mathcal{X}}$ and $\pi_{\mathcal{B}'_\mathcal{X}}$ such that $\mathbb{E}[1_{\Psi_\mathcal{X},v}] \leq \alpha(1-\varepsilon)^{d(w_\mathcal{X},v)}$ for every vertex $v \in R_\mathcal{X}$, where $\alpha > 0$ and $0 < \varepsilon < 1$. Instead of summing over the vertices in $R_\mathcal{X}$, we sum over all vertices in $V_\mathcal{G}$ except for $w_\mathcal{X}$. The sum can be written as a sum over all distances d from $w_\mathcal{X}$. Thus,

$$\begin{aligned}\sum_{v\in R_\mathcal{X}}\mathbb{E}[1_{\Psi_\mathcal{X},v}] \leq \sum_{v\in R_\mathcal{X}}\alpha(1-\varepsilon)^{d(w_\mathcal{X},v)} &\leq \sum_{v\in V_\mathcal{G}\setminus\{w_\mathcal{X}\}}\alpha(1-\varepsilon)^{d(w_\mathcal{X},v)} \\ &= \sum_{d\geq 1} n_d(w_\mathcal{X})\alpha(1-\varepsilon)^d \leq K,\end{aligned}$$

where K is a constant since the graph $\mathcal{G}$ is uniformly sub-exponential (see Definition 2.18) and hence $n_d(w_\mathcal{X})$ grows slower than exponentially. □

Now when the implication from exponential decay of edge-boundary pairs to strong spatial mixing and K-coupling cover has been established, it remains to show for what graphs $\mathcal{G}$, number of colours q and values λ there is exponential decay. In the next chapter we define a coupling Ψ of π_{B_X} and $\pi_{B'_X}$ for any edge-boundary pair X. We use this coupling in subsequent chapters when showing exponential decay for different lattices and parameters q and λ. The coupling Ψ is defined recursively and was first introduced by Goldberg, Martin and Paterson in [GMP05].

Chapter 6

Recursive coupling

In this chapter we describe how to construct a coupling Ψ of π_{B_X} and $\pi_{B'_X}$ for edge-boundary pairs X. We would like to show that there are two constants $\alpha > 0$ and $0 < \varepsilon < 1$ such that for any edge-boundary pair X, the quantity $\mathbb{E}[1_{\Psi,v}] \leq \alpha(1-\varepsilon)^{d(e_X,v)}$, where $\mathbb{E}[1_{\Psi,v}]$ is the probability that a vertex $v \in R_X$ receive different colours in two colourings of R_X drawn from Ψ. This would imply, by using Lemma 5.3 followed by Lemma 5.4, that the system has strong spatial mixing. Furthermore, from Lemma 5.3 followed by Lemma 5.5 we have that the system has a K-coupling cover, provided the graph $\mathcal{G}$ is uniformly sub-exponential. Applying Theorem 2.15 then implies that the Glauber dynamics is rapidly mixing (provided $\lambda > 0$ or $q \geq \Delta + 2$). We use the construction of Ψ to prove Theorem 3.1.

Goldberg, Martin and Paterson invented in [GMP05] a method for recursively constructing couplings Ψ of π_{B_X} and $\pi_{B'_X}$ for edge-boundary pairs X. We use their technique in this thesis. Couplings that are constructed recursively have been used in the past for mixing on trees, but was extended in [GMP05] to non-tree-like graphs such as lattices. From the construction of Ψ we derive an upper bound on the quantity $\mathbb{E}[1_{\Psi,v}]$ which is hopefully decreasing exponentially as above. Whether there is exponential decay or not depends on the spin system. If we fail to show exponential decay of $\mathbb{E}[1_{\Psi,v}]$ with the general technique, then it could help to consider the geometry of the lattice in order to derive an upper bound on $\mathbb{E}[1_{\Psi,v}]$ that is decreasing exponentially. The idea of making use of the geometry has been demonstrated in [GMP05] and has led to improved results such as strong spatial mixing on the triangular lattice and $\mathbb{Z}^3$ with parameters $\lambda = 0$ and $q = 10$. In this chapter we present a framework in which the geometry of a lattice can be exploited in a systematic way, preferably with the help of a computer. We define the notion of an $(\mathcal{A}, \mathcal{F})$-set and describe how it relates to exponential decay of $\mathbb{E}[1_{\Psi,v}]$. In subsequent chapters we will construct $(\mathcal{A}, \mathcal{F})$-sets for different lattices and parameters q and λ, and derive mixing results that improve the bounds given by the general theorems for mixing.

The chapter is organised as follows. In Section 6.1 we define the recursive coupling Ψ of π_{B_X} and $\pi_{B'_X}$ as a tree, denoted T_X. In Section 6.2 we show how the tree T_X is related to the quantity $\mathbb{E}[1_{\Psi,v}]$, and in Section 6.3 we examine under what conditions there is exponential decay of $\mathbb{E}[1_{\Psi,v}]$. This is where we state the proof of Theorem 3.1. In Section 6.4 we define the notion of an $(\mathcal{A},\mathcal{F})$-set and explain how it is used to improve mixing bounds for specific lattices. In Section 6.5 we introduce the notion of an extended region which will be useful in subsequent chapters when constructing $(\mathcal{A},\mathcal{F})$-sets.

6.1 The recursive coupling tree

Consider any edge-boundary pair X and let $E_X = \{e_1,\dots,e_{n_X}\}$ be an ordered set of the edges which connect v_X to another vertex in R_X. If $\mathcal{G}$ is planar, let $e_1,\dots,e_{n_X}$ be the clockwise ordering of the edges in E_X around v_X such that the discrepancy edge e_X (i.e., edge (w_X,v_X)) appears between edge e_{n_X} and e_1. This guarantees that e_i and e_j are adjacent only if i and j differ by 1. Recall the definition of adjacent edges on Page 52. If $\mathcal{G}$ is not planar, take any ordering of the edges in E_X. For each $i \in \{1,\dots,n_X\}$ and two colours $c,c' \in Q$, where $c \neq c'$, let $X_i(c,c')$ be the edge-boundary pair consisting of

- The region $R_{X_i(c,c')} = R_X \backslash \{v_X\}$,
- the distinguished boundary edge $e_{X_i(c,c')} = e_i$, and
- the pair $(B_{X_i(c,c')}, B'_{X_i(c,c')})$ of q_0-colourings of $\mathcal{E}R_{X_i(c,c')}$ such that both colourings are identical to B_X (and hence B'_X) on all edges in $\mathcal{E}R_{X_i(c,c')} \backslash E_X$. The colours of the edges in E_X are assigned as follows.
  - $B_{X_i(c,c')}(e_j) = B'_{X_i(c,c')}(e_j) = c'$ for $j = 1,\dots,i-1$,
  - $B_{X_i(c,c')}(e_i) = c'$ and $B'_{X_i(c,c')}(e_i) = c$, and
  - $B_{X_i(c,c')}(e_j) = B'_{X_i(c,c')}(e_j) = c$ for $j = i+1,\dots,n_X$.

Note that $X_i(c,c')$ satisfies all conditions of being an edge-boundary pair, also when $\mathcal{G}$ is planar. The edge-boundary pairs $X_i(c,c')$ represent boundary pairs that are constructed from X after shrinking the region by removing vertex v_X and keeping the colours of all boundary edges. The new boundary edges that are incident to v_X (which is now a boundary vertex) are given colours c and c' similarly to the sequence of edge colourings derived in Section 5.2 when breaking the discrepancy at a vertex into discrepancies at edges around it. Figure 6.1 illustrates the idea. Let $\Psi_i(c,c')$ be a coupling of $\pi_{B_{X_i(c,c')}}$ and $\pi_{B'_{X_i(c,c')}}$. The coupling $\Psi_{c,c'}$ of $\pi_{B_{X_1(c,c')}}$ and $\pi_{B'_{X_{n_X}(c,c')}}$ is constructed

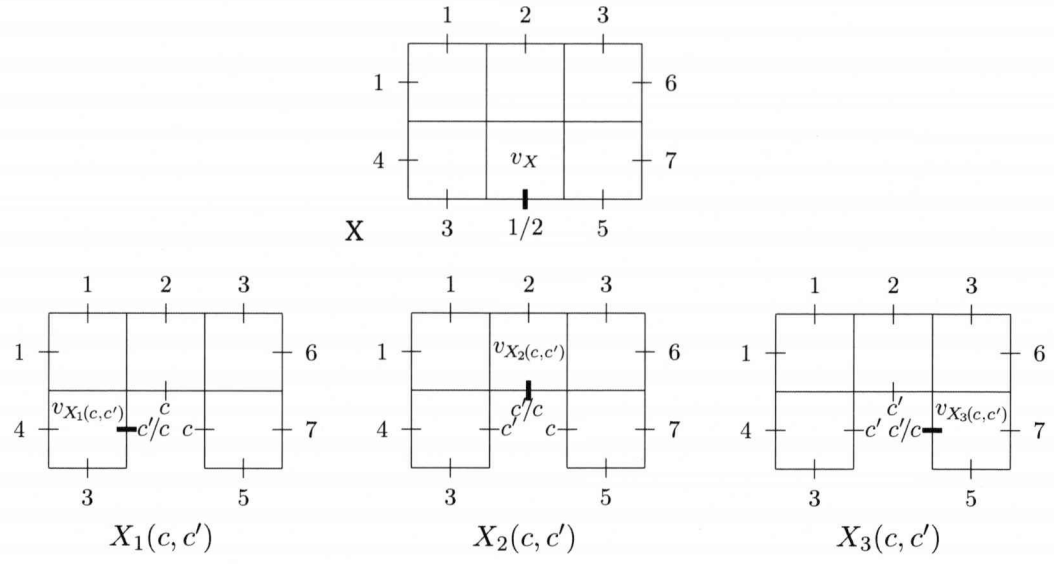


Figure 6.1. An edge-boundary pair X on the square lattice, and the three edge-boundary pairs $X_1(c,c')$, $X_2(c,c')$ and $X_3(c,c')$ that are constructed recursively in the tree T_X.

by composing the couplings $\Psi_1(c,c'),\dots,\Psi_{n_X}(c,c')$ in the same way as the couplings $\Psi_1,\dots,\Psi_k$ in Section 5.2 were composed. That is, to choose a pair (σ_0,σ_{n_X}) from $\Psi_{c,c'}$, first choose the pair (σ_0,σ_1) from $\Psi_1(c,c')$. Suppose that $\sigma_0 = x_0$ and $\sigma_1 = x_1$. Then choose the pair (σ_1,σ_2) from the conditional distribution $\Psi_2(c,c')$, conditioned on $\sigma_1 = x_1$. Suppose that $\sigma_2 = x_2$. Then choose the pair (σ_2,σ_3) from the conditional distribution $\Psi_3(c,c')$, conditioned on $\sigma_2 = x_2$, and so on. Hence, σ_0 is drawn from $\pi_{B_{X_1(c,c')}}$ and σ_{n_X} is drawn from $\pi_{B'_{X_{n_X}(c,c')}}$.

Recall from Section 5.2 that for a vertex $v \in R_X$ and a coupling Ψ of π_{B_X} and $\pi_{B'_X}$, $1_{\Psi,v}$ is the indicator random variable for the event that the colour of v differs in two colourings drawn from Ψ. We define $\Psi_X^{\min}$ to be some coupling of π_{B_X} and $\pi_{B'_X}$ minimising $\mathbb{E}[1_{\Psi,v_X}]$ over all couplings Ψ. That is, if (C,C') is a pair of colourings of R_X drawn from $\Psi_X^{\min}$, then the probability that $C(v_X) \neq C'(v_X)$ is minimised. For every pair of colours $c,c' \in Q$, let $p_X^{\min}(c,c')$ be the probability that $C(v_X) = c$ and $C'(v_X) = c'$.

In order to construct a coupling Ψ of π_{B_X} and $\pi_{B'_X}$ we define an edge labelled tree T_X. After T_X has been defined we show how certain properties of the tree can be related to $\mathbb{E}[1_{\Psi,v}]$, the probability that the colour of v differs in the two colourings drawn from Ψ. The tree T_X is defined recursively as follows. It might be helpful to look at Figure 6.2 which illustrates the tree T_X for the edge-boundary pair X from Figure 6.1. Start with a node r which will be the root of T_X (the edge-boundary pair X is illustrated in the root of T_X in Figure 6.2). For every pair of colours $c,c' \in Q$, where $c \neq c'$, add an edge labelled $(p_X^{\min}(c,c'),v_X)$ from r to a new node $r_{c,c'}$. If E_X is non-empty then recursively construct the trees $T_{X_i(c,c')}$, $i = 1,\dots,n_X$, and join them to

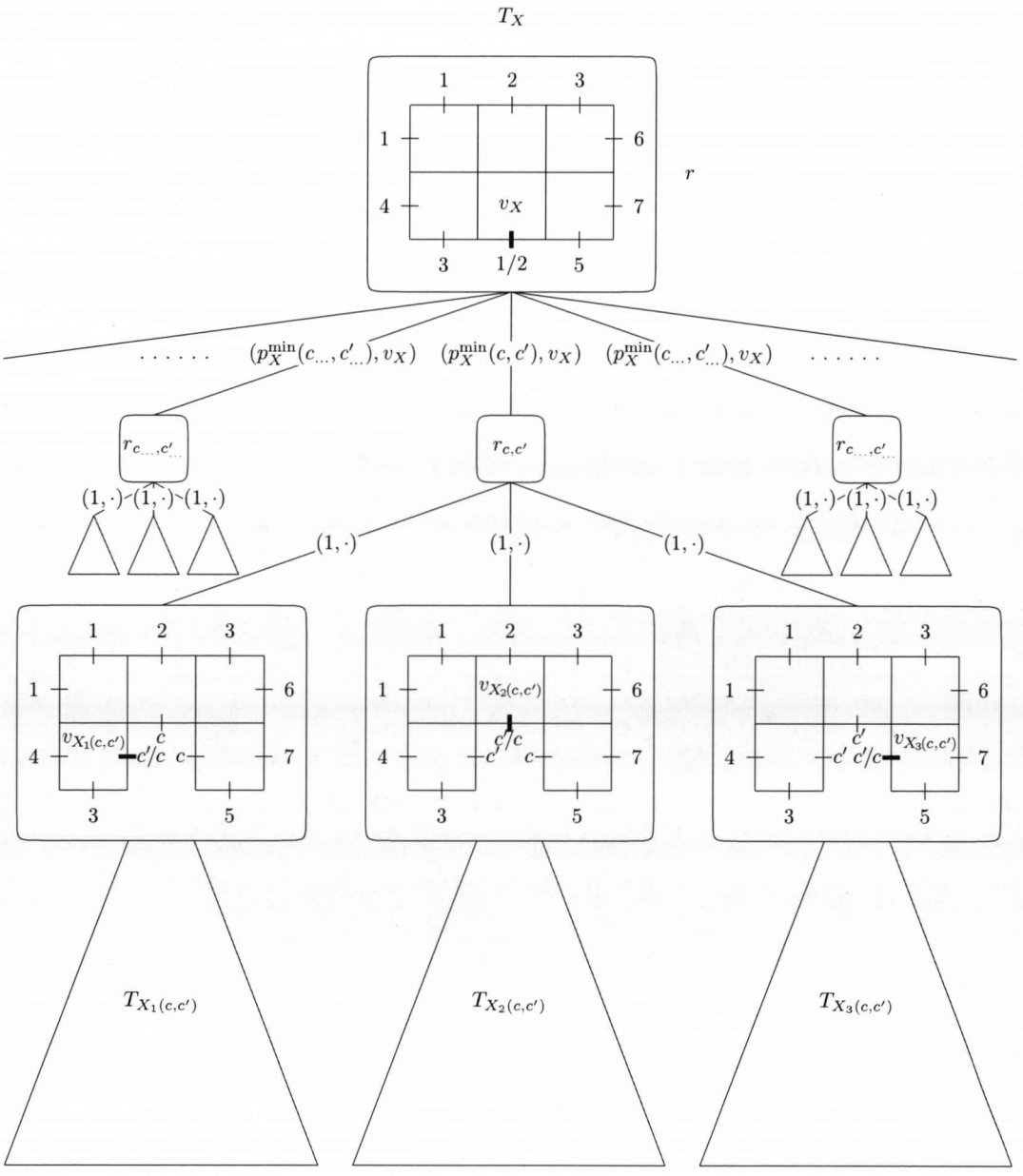


Figure 6.2. The recursive coupling tree T_X.

T_X by adding an edge with the label $(1,\cdot)$ from $r_{c,c'}$ to the root of each tree $T_{X_i(c,c')}$. If E_X is empty then $r_{c,c'}$ is a leaf. That completes the construction of T_X. In Figure 6.2, the edge-boundary pair $X_i(c,c')$ is illustrated in the root of $T_{X_i(c,c')}$.

6.1.1 The purpose of the recursive coupling tree

Before formally describing how properties of the recursive coupling tree T_X are connected to the quantity $\mathbb{E}[1_{\Psi,v}]$, we give a slightly less formal explanation of the connection.

Suppose that X is the edge-boundary pair from Figure 6.1. A coupling Ψ of π_{B_X} and $\pi_{B'_X}$ is defined as follows (a formal definition is given in the proof of Lemma 6.1 in the next section). We use the coupling $\Psi_X^{\min}$ to choose the colours of the vertex v_X. Hence, the probability that v_X receives the colours c and c', respectively, in two colourings drawn from Ψ is $p_X^{\min}(c,c')$. In order to complete the definition of Ψ, choose two colourings of the remaining vertices of R_X by using the the coupling $\Psi_{c,c'}$ in the previous section. As we have seen, the coupling $\Psi_{c,c'}$ can be defined as the composition of the couplings $\Psi_i(c,c')$. We use this definition of $\Psi_{c,c'}$ in order to define Ψ. Each of the couplings $\Psi_i(c,c')$ is defined recursively in the same manner as Ψ.

Let $v \neq v_X$ be a vertex of the region R_X. Let (C,C') denote the random variable corresponding to a pair of colourings of R_X drawn from Ψ. Conditioned on the event that $C(v_X) = c$ and $C'(v_X) = c'$, let $p_{c,c'}(v)$ be the probability that $C(v) \neq C'(v)$. From the definition of Ψ we have that $p_{c,c'}(v)$ is the probability that v receives different colours in a pair of colourings drawn from $\Psi_{c,c'}$. From the definition of $\Psi_{c,c'}$ as the composition of the couplings $\Psi_i(c,c')$, it follows that if $C(v) \neq C'(v)$ then at least one of the couplings $\Psi_i(c,c')$ must have returned two colourings in which the colour of v differs. Therefore, $p_{c,c'}(v) \leq \hat{p}_{c,c'}(v)$, where $\hat{p}_{c,c'}(v)$ is the sum, over i, of the probabilities that v differs in colourings drawn from $\Psi_i(c,c')$. We now have that

$$\mathbb{E}[1_{\Psi,v}] \leq \sum_{c,c'} p_X^{\min}(c,c') \cdot \hat{p}_{c,c'}(v), \tag{6.1}$$

where the sum is over all possible choices of colours c and c' on v_X. The recursive construction of the coupling Ψ resembles the recursive construction of the tree T_X. The sum in Equation (6.1) can be extracted from the tree T_X in the following way.

Identify edges in T_X for which the last component of the label is v. Figure 6.3 illustrates a path from the root of T_X down to an edge e for which the last component of its label is v. The product of the first components of the edge labels along this path is

$$p_X^{\min}(c,c') \cdot p_{X'}^{\min}(b,b') \cdot p_{X''}^{\min}(a,a'),$$

where c, c', b, b', a and a' are colours in Q, and X' and X'' are edge-boundary pairs

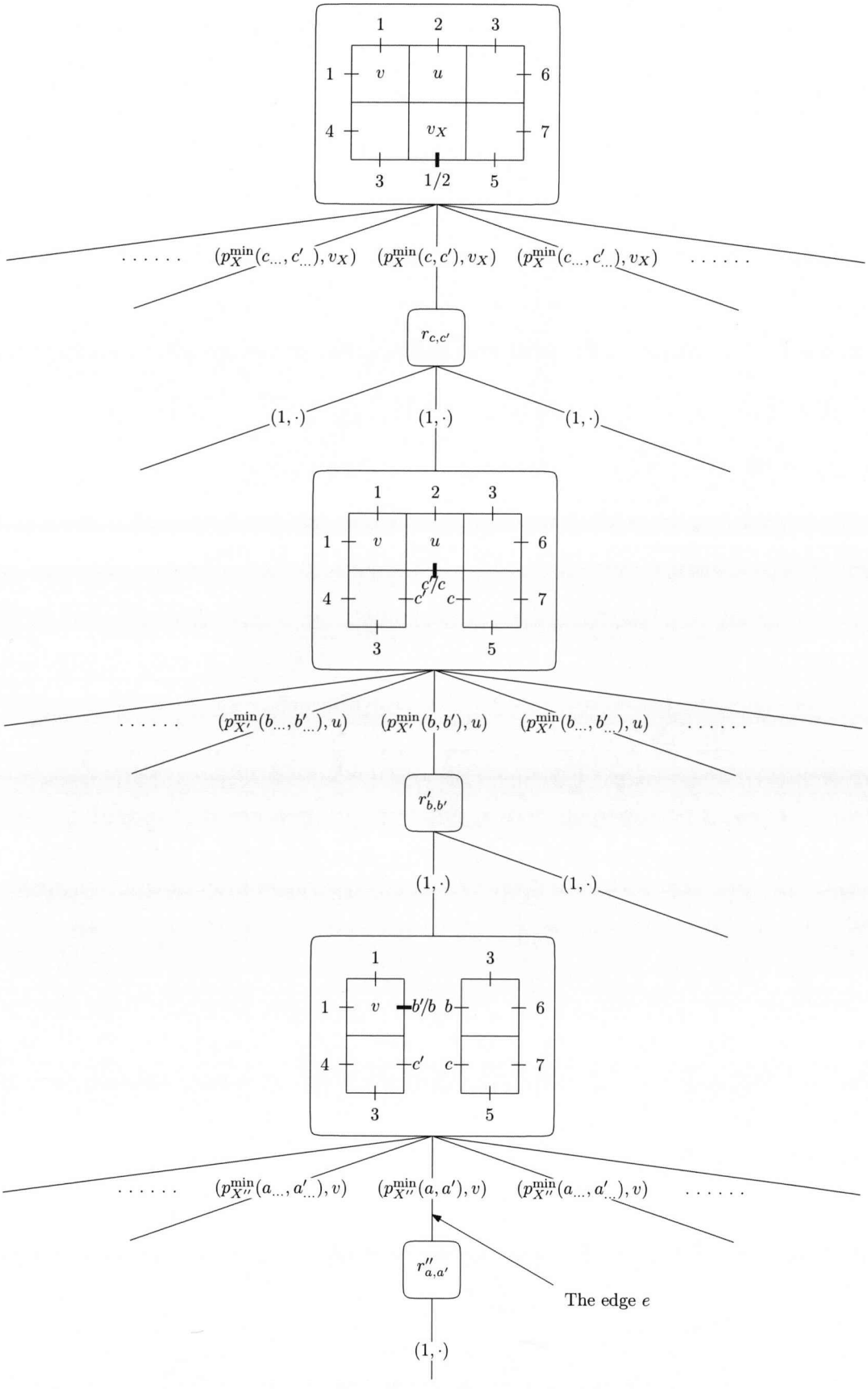


Figure 6.3. A path in the recursive coupling tree T_X

defined recursively in T_X. Note that edges that are labelled $(1, \cdot)$ are given this label because the first component 1 will enable us to correctly define the product of edge weights (using the first component of the label) over the edges in a path. We will refer to the product of the edge weights along a path from the root down to an edge e as the *likelihood* of e. The word likelihood is chosen because it is a product of probabilities which somewhat reflects the probability of choosing a specific colouring of a particular vertex-path in R_X. The *cost* of the vertex v will be defined as the sum of the likelihoods over all edges e for which the second component is v. The cost of v is an upper bound on $\mathbb{E}[1_{\Psi,v}]$. In fact, the cost of v is the sum in Equation (6.1). The next section will explain this formally.

6.1.2 Properties of the recursive coupling tree

We introduce some notation, taken from [GMP05] in order to link properties of the tree T_X to the probability $\mathbb{E}[1_{\Psi,v}]$. We say that an edge e of T_X is *degenerate* if the second component of its label is "$\cdot$". Note that if we follow a path from the root of T_X to any of its leaves then every second edge is a degenerate edge. For edges e and e' of T_X, we write $e \to e'$ to denote the fact that e is an ancestor of e'. That is, either $e = e'$, or e is a proper ancestor of e'. Define the *level* of an edge e of T_X to be the number of non-degenerate edges on the path from the root down to, and including, e. Suppose that e is an edge of T_X with label (p, v). We say that the *weight* $w(e)$ of edge e is p. Also the *name* $n(e)$ of edge e is v. The *likelihood* $l(e)$ of e is $\prod_{e':e'\to e} w(e')$. The *cost* $\gamma(v, T_X)$ of a vertex $v \in R_X$ is $\sum_{e:n(e)=v} l(e)$. If the region R_X is not connected and vertex v_X and a vertex $v \in R_X$ belong to different connected components, then there will be no edge with name v in T_X and we define $\gamma(v, T_X) = 0$. We have the following lemma ([GMP05, Lemma 12]). We provide the proof for completeness.

Lemma 6.1. *For every edge-boundary pair X there exists a coupling Ψ of π_{B_X} and $\pi_{B'_X}$ such that $\mathbb{E}[1_{\Psi,v}] \leq \gamma(v, T_X)$ for all $v \in R_X$.*

Proof. Let X be an edge-boundary pair. The coupling Ψ is constructed recursively in the same manner as the tree T_X, where at each stage, the discrepancy at a given vertex is broken to discrepancies at single edges, so at every stage of the recursion we only need to consider a pair of colourings with a discrepancy at a single edge (i.e., an edge-boundary pair).

We will construct Ψ so that if R_X is not connected and v belongs to a connected component R that does not contain v_X, then we always colour R with the same colouring in both colourings drawn from Ψ, and hence $\mathbb{E}[1_{\Psi,v}] = 0$. Now consider a region R_X that is connected. Let (C, C') denote the random variable corresponding to a pair of colourings drawn from the coupling Ψ that we are constructing. If R_X consists only

of vertex v_X then we define $\Psi = \Psi_X^{\min}$. If R_X consists of more vertices then we use $\Psi_X^{\min}$ to couple the colouring of vertex v_X. Conditioning on this, we recursively construct a different coupling in order to choose colourings of the other vertices in R_X. We assign $C(v_X) = c$ and $C'(v_X) = c'$ with probability $p_X^{\min}(c, c')$.

Let R be the region $R_X \backslash \{v_X\}$ and let B_c be the edge colouring of $\mathcal{E}R$ that is identical to the colouring B_X on edges $\mathcal{E}R \cap \mathcal{E}R_X$ and which assigns colour c on edges in $\mathcal{E}R$ that are incident to v_X. Let $B_{c'}$ be the edge colouring of $\mathcal{E}R$ that is identical to B_c but assigns colour c' on the edges in $\mathcal{E}R$ that are incident to v_X. Let $\Psi_{c,c'}$ be a coupling of π_{B_c} and $\pi_{B_{c'}}$. Thus, in order to construct the coupling Ψ, we first choose a pair (c, c') of colours on v_X by using the coupling $\Psi_X^{\min}$. Once the colours c and c' are selected, we complete the construction of Ψ by choosing two colourings from $\Psi_{c,c'}$ in order to colour the rest of the vertices in R_X. How we construct $\Psi_{c,c'}$ is explained next.

If $c = c'$ then let $\Psi_{c,c'}$ be the perfect coupling where the two colourings drawn from $\Psi_{c,c'}$ are identical. If $c \neq c'$ then the coupling $\Psi_{c,c'}$ is constructed as follows. As previously described, let $E_X = \{e_1, \ldots, e_{n_X}\}$ be the ordered set of edges which connect v_X to another vertex in R_X. Let $X_i(c, c')$, $i = 1, \ldots, n_X$, be the edge-boundary pairs defined previously. We now have $\pi_{B_c} = \pi_{B_{X_1(c,c')}}$ and $\pi_{B_{c'}} = \pi_{B'_{X_{n_X}(c,c')}}$, and hence $\Psi_{c,c'}$ is a coupling of $\pi_{B_{X_1(c,c')}}$ and $\pi_{B'_{X_{n_X}(c,c')}}$. As defined earlier in this section, let $\Psi_i(c, c')$ be a coupling of $\pi_{B_{X_i(c,c')}}$ and $\pi_{B'_{X_i(c,c')}}$. We construct $\Psi_{c,c'}$ by composing the couplings $\Psi_1(c, c'), \ldots, \Psi_{n_X}(c, c')$ as explained in the beginning of this section.

We will now show that $\mathbb{E}[1_{\Psi,v}] \leq \gamma(v, T_X)$ for all $v \in R_X$. The proof is by induction on $|R_X|$. If $v = v_X$ then $\mathbb{E}[1_{\Psi,v}] = \gamma(v, T_X)$ by the construction of Ψ and T_X. This handles the base case when $|R_X| = 1$. Suppose $v \neq v_X$ and $|R_X| > 1$. Then

$$\begin{aligned}
\mathbb{E}[1_{\Psi,v}] &= \sum_{\substack{c,c' \in Q,\\ c \neq c'}} p_X^{\min}(c, c') \mathbb{E}[1_{\Psi_{c,c'},v}] \leq \sum_{\substack{c,c' \in Q,\\ c \neq c'}} p_X^{\min}(c, c') \sum_{i=1}^{n_X} \mathbb{E}[1_{\Psi_i(c,c'),v}] \\
&\leq \sum_{\substack{c,c' \in Q,\\ c \neq c'}} p_X^{\min}(c, c') \sum_{i=1}^{n_X} \gamma(v, T_{X_i(c,c')}) = \gamma(v, T_X),
\end{aligned}$$

where the first inequality follows from the fact that $\Psi_{c,c'}$ is constructed as a composition of the couplings $\Psi_i(c, c')$, and the second inequality uses the inductive hypothesis. □

For any edge-boundary pair X and any integer $d \geq 1$, let $E_d(X)$ denote the set of level-d edges in T_X, and let $\Gamma^d(X) = \sum_{e \in E_d(X)} l(e)$. We will often refer to the following equivalent definition of $\Gamma^d(X)$. For $d = 1$ we have

$$\Gamma^1(X) = \sum_{\substack{c,c' \in Q,\\ c \neq c'}} p_X^{\min}(c, c').$$

For $d > 1$ we have

$$\Gamma^d(X) = \sum_{\substack{c,c' \in Q, \\ c \neq c'}} p_X^{\min}(c, c') \sum_{i=1}^{n_X} \Gamma^{d-1}(X_i(c, c')). \tag{6.2}$$

The quantity $\Gamma^d(X)$ has a key role when we show that a spin system has a strong spatial mixing and a K-coupling cover. If $\Gamma^d(X)$ is decreasing exponentially with d, then there exists a coupling Ψ of π_{B_X} and $\pi_{B'_X}$ such that $\mathbb{E}[1_{\Psi,v}]$ is decreasing exponentially with the distance between the discrepancy edge e_X and a vertex $v \in R_X$. This fact is stated in the following Lemma.

Lemma 6.2. *Suppose that there are constants $\alpha > 0$ and $0 < \varepsilon < 1$ such that, for every edge-boundary pair X and every non-negative integer d, $\Gamma^d(X) \leq \alpha(1-\varepsilon)^d$. Then there is a constant $\alpha' > 0$ such that for every edge-boundary pair X there is a coupling Ψ of π_{B_X} and $\pi_{B'_X}$ such that for every $v \in R_X$,*

$$\mathbb{E}[1_{\Psi,v}] \leq \alpha'(1-\varepsilon)^{d(e_X,v)}.$$

Proof. Suppose that there are constants $\alpha > 0$ and $0 < \varepsilon < 1$ such that, for every edge-boundary pair X and every non-negative integer d, $\Gamma^d(X) \leq \alpha(1-\varepsilon)^d$. Let X be an edge-boundary pair. From Lemma 6.1 there is a coupling Ψ of π_{B_X} and $\pi_{B'_X}$ such that, for every $v \in R_X$, we have $\mathbb{E}[1_{\Psi,v}] \leq \gamma(v, T_X)$. Let Ψ be such a coupling. Let v be a vertex in R_X. Note that for $d < d(e_X, v)$ there is no level-d edge e in T_X with name $n(e) = v$. Thus, in order to upper-bound $\gamma(v, T_X)$ we sum the likelihoods $l(e)$ over all edges e at level $d \geq d(e_X, v)$. Thus,

$$\begin{aligned}
\mathbb{E}[1_{\Psi,v}] \leq \gamma(v, T_X) &= \sum_{e:n(e)=v} l(e) \leq \sum_{d \geq d(e_X,v)} \sum_{e \in E_d(X)} l(e) \\
&= \sum_{d \geq d(e_X,v)} \Gamma^d(X) \leq \sum_{d \geq d(e_X,v)} \alpha(1-\varepsilon)^d = \alpha'(1-\varepsilon)^{d(e_X,v)},
\end{aligned}$$

where $\alpha' = \alpha/\varepsilon$. □

Thus, if $\Gamma^d(X) \leq \alpha(1-\varepsilon)^d$ then it follows that the system has strong spatial mixing and that the Glauber dynamics is rapidly mixing (provided $\mathcal{G}$ is uniformly sub-exponential and either $\lambda > 0$ or $q \geq \Delta + 2$). We state this in the following lemma, which we will use in subsequent chapters when showing strong spatial mixing and rapid mixing.

Lemma 6.3. *Suppose that there are constants $\alpha > 0$ and $0 < \varepsilon < 1$ such that, for every edge-boundary pair X and every non-negative integer d, $\Gamma^d(X) \leq \alpha(1-\varepsilon)^d$. Then*

(1) the spin system has strong spatial mixing, and

(2) if $\mathcal{G}$ is uniformly sub-exponential, and $\lambda > 0$ or $q \geq \Delta+2$, then the Glauber dynamics is rapidly mixing for any region R and any q_0-colouring $\mathcal{B}$ of ∂R. More precisely, the mixing time $\tau(\delta) \in O(n^2 + n \log \frac{1}{\delta})$, where n is the number of vertices in R.

Proof. To prove (1), we use Lemma 6.2, 5.3 and 5.4. To prove (2), we use Lemma 6.2, 5.3 and 5.5 to show that the spin system has a K-coupling cover. The mixing time of the Glauber dynamics then follows from Theorem 2.15. □

6.2 Bounding the costs in the recursive coupling tree

Consider any edge-boundary pair X. A key ingredient from the construction of T_X which affects the cost $\gamma(v, T_X)$ is the quantity $\mathbb{E}[1_{\Psi_X^{\min}, v_X}]$ which we denote $\nu(X)$. Thus,

$$\nu(X) = \mathbb{E}[1_{\Psi_X^{\min}, v_X}] = \sum_{\substack{c,c' \in Q, \\ c \neq c'}} p_X^{\min}(c, c').$$

If we obtain upper bounds on $\nu(X)$, then these upper bounds can be used to upper-bound the costs in the tree T_X, from which we derive an upper bound on $\Gamma^d(X)$. Goldberg, Martin and Paterson [GMP05] observed that $\nu(X)$ is upper-bounded by a value which they denoted $\mu(X)$. The value $\mu(X)$ is defined in [GMP05] only for $\lambda = 0$. It was demonstrated in [GMP05] (for $\lambda = 0$) that $\mu(X)$ is upper-bounded by $\mu(X')$, where X' is an edge-boundary pair derived from X such that the region $R_{X'}$ is a subset of R_X. In particular, if $R_{X'}$ consists of only vertex $v_{X'}$, then an upper bound on the value $\mu(X')$ can be computed by hand and would provide an upper bound on $\nu(X)$ for all edge-boundary pairs X. This idea has led to improved mixing results on general graphs (Theorem 3.1). However, in order to obtain better results, larger regions have to be considered, which, as we will see, involves computer-assisted computations. We give the definition of $\mu(X)$ next.

Recall from Section 5.2 that $\text{mon}_{\sigma, B_X}(E)$ is the number of monochromatic edges in $E \subseteq E(R_X)$ under a colouring σ of R_X and boundary colouring B_X. For each colour $i \in Q$, let $\Omega^+_{R_X}(v_X = i)$ denote the set of colourings in $\Omega^+_{R_X}$ that assign colour i to vertex v_X. Let ω_i be the *weight* of these colourings, ignoring edge e_X. That is,

$$\omega_i = \sum_{\sigma \in \Omega^+_{R_X}(v_X = i)} \lambda^{\text{mon}_{\sigma, B_X}(E(R_X) - \{e_X\})}.$$

For any two colours $c, c' \in Q$, let

$$\hat{\omega}_{c,c'} = \sum_{i \in Q \setminus \{c,c'\}} \omega_i.$$

We now define

$$\mu_{c,c'}(X) = \frac{(1-\lambda)\omega_c}{(1+\lambda)\omega_c + \hat{\omega}_{c,c'}}$$

and

$$\mu(X) = \max[\mu_{B_X(e_X),B'_X(e_X)}(X),\ \mu_{B'_X(e_X),B_X(e_X)}(X)].$$

The following lemma enables us to use $\mu(X)$ as an upper bound for $\nu(X)$.

Lemma 6.4. *For every edge-boundary pair X, $\nu(X) \leq \mu(X)$.*

Proof. Let $\pi_{B_X,v_X}(i)$ and $\pi_{B'_X,v_X}(i)$ be the probability that vertex v_X receives colour i in a colouring drawn from π_{B_X} and $\pi_{B'_X}$, respectively. Without loss of generality (to simplify notation) suppose $B_X(e_X) = 1$, $B'_X(e_X) = 2$, and $\omega_1 \geq \omega_2$. Since the weight ω_i does not include edge e_X we have

$$\pi_{B_X,v_X}(i) = \begin{cases} \frac{\lambda\omega_1}{\lambda\omega_1+\omega_2+\hat{\omega}_{1,2}}, & \text{if } i = 1; \\ \frac{\omega_i}{\lambda\omega_1+\omega_2+\hat{\omega}_{1,2}}, & \text{otherwise,} \end{cases}$$

$$\pi_{B'_X,v_X}(i) = \begin{cases} \frac{\lambda\omega_2}{\omega_1+\lambda\omega_2+\hat{\omega}_{1,2}}, & \text{if } i = 2; \\ \frac{\omega_i}{\omega_1+\lambda\omega_2+\hat{\omega}_{1,2}}, & \text{otherwise.} \end{cases}$$

In order to compare the probabilities π_{B_X,v_X} and $\pi_{B'_X,v_X}$, we scale them with their common divisor $Z = (\lambda\omega_1 + \omega_2 + \hat{\omega}_{1,2})(\omega_1 + \lambda\omega_2 + \hat{\omega}_{1,2})$. Thus,

$$\pi_{B_X,v_X}(i) = \begin{cases} \lambda\omega_1(\omega_1 + \lambda\omega_2 + \hat{\omega}_{1,2})/Z, & \text{if } i = 1; \\ \omega_i(\omega_1 + \lambda\omega_2 + \hat{\omega}_{1,2})/Z, & \text{otherwise,} \end{cases}$$

$$\pi_{B'_X,v_X}(i) = \begin{cases} \lambda\omega_2(\lambda\omega_1 + \omega_2 + \hat{\omega}_{1,2})/Z, & \text{if } i = 2; \\ \omega_i(\lambda\omega_1 + \omega_2 + \hat{\omega}_{1,2})/Z, & \text{otherwise.} \end{cases}$$

Since $\omega_1 \geq \omega_2$ and $\lambda \in [0,1]$, it is now straightforward to verify that the scaled numerator in the expression for $\pi_{B_X,v_X}(1)$ is less than or equal to the scaled numerator in $\pi_{B'_X,v_X}(1)$. For $i = 2,\ldots,q$, the opposite is true. That is, the scaled numerator in the expression for $\pi_{B_X,v_X}(i)$ is greater than or equal to the scaled numerator in $\pi_{B'_X,v_X}(i)$. The coupling $\Psi_X^{\min}$ of π_{B_X} and $\pi_{B'_X}$ is therefore constructed such that a disagreement at vertex v_X occurs only when the sample from $\pi_{B'}$ assigns colour 1 to v_X. If the sample from $\pi_{B'}$ assigns colour $i \geq 2$ to v_X then we choose a sample from π_B that also assigns colour i to v_X. Thus, the probability of having a disagreement at v_X is

$\pi_{B'_X,v_X}(1) - \pi_{B_X,v_X}(1)$. We have

$$\begin{aligned}\nu(X) &= \pi_{B'_X,v_X}(1) - \pi_{B_X,v_X}(1) = \frac{\omega_1(\lambda\omega_1 + \omega_2 + \hat{\omega}_{1,2}) - \lambda\omega_1(\omega_1 + \lambda\omega_2 + \hat{\omega}_{1,2})}{(\lambda\omega_1 + \omega_2 + \hat{\omega}_{1,2})(\omega_1 + \lambda\omega_2 + \hat{\omega}_{1,2})} \\ &= \frac{\omega_1(1-\lambda)(\omega_2 + \lambda\omega_2 + \hat{\omega}_{1,2})}{(\lambda\omega_1 + \omega_2 + \hat{\omega}_{1,2})(\omega_1 + \lambda\omega_2 + \hat{\omega}_{1,2})} \leq \frac{(1-\lambda)\omega_1}{(1+\lambda)\omega_1 + \hat{\omega}_{1,2}} = \mu_{1,2}(X) \\ &\leq \mu(X)\end{aligned}$$

which proves the lemma. □

The next lemma says that an upper bound on $\mu_{c,c'}(X)$ can be obtained by considering another edge-boundary pair X', whose region $R_{X'}$ is a subregion of R_X. If $\mu_{c,c'}(X')$ is maximised over possible boundary colourings of $R_{X'}$ then this value is an upper bound on $\mu_{c,c'}(X)$.

Lemma 6.5. *Suppose that X is an edge-boundary pair and suppose c and c' are the two colours of edge e_X in B_X and B'_X, in any order. Let R' be any subset of R_X which includes v_X. Let S be the set of edge-boundary pairs X' such that $R_{X'} = R'$, the distinguished edge $e_{X'} = e_X$, and the boundary colouring $B_{X'}$ is identical to B_X, and $B'_{X'}$ is identical to B'_X on the edges $\mathcal{E}R_X \cap \mathcal{E}R'$. Then $\mu_{c,c'}(X) \leq \max_{X' \in S} \mu_{c,c'}(X')$.*

Proof. Without loss of generality, suppose $c = B_X(e_X)$ and $c' = B'_X(e_X)$. Let $H = R_X \backslash R'$. For colour $i \in Q$ and colouring $\theta \in \Omega^+_H$, let $\Omega^+_{R_X,i,\theta}$ be the set of colourings $\sigma \in \Omega^+_{R_X}$ with $\sigma(v_X) = i$ and $\sigma(H) = \theta$. Let

$$\omega_{i,\theta} = \sum_{\sigma \in \Omega^+_{R_X,i,\theta}} \lambda^{\mathrm{mon}_{\sigma,B_X}(E(R_X) - \{e_X\})}$$

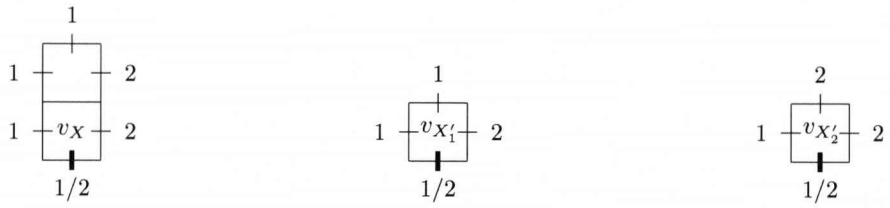
and let

$$\hat{\omega}_{c,c',\theta} = \sum_{i \in Q \backslash \{c,c'\}} \omega_{i,\theta}.$$

Then

$$\begin{aligned}\mu_{c,c'}(X) &= \frac{(1-\lambda)\omega_c}{(1+\lambda)\omega_c + \hat{\omega}_{c,c'}} = \frac{\sum_{\theta \in \Omega^+_H}(1-\lambda)\omega_{c,\theta}}{\sum_{\theta \in \Omega^+_H}((1+\lambda)\omega_{c,\theta} + \hat{\omega}_{c,c',\theta})} \\ &\leq \max_{\theta \in \Omega^+_H} \frac{(1-\lambda)\omega_{c,\theta}}{(1+\lambda)\omega_{c,\theta} + \hat{\omega}_{c,c',\theta}} \leq \max_{X' \in S} \mu_{c,c'}(X').\end{aligned}$$

To see the last inequality, take any $\theta \in \Omega^+_H$ and construct the edge-boundary pair X' with the following parameters: $R_{X'} = R'$, $B_{X'}$ is identical to B_X, and $B'_{X'}$ is identical to B'_X on the edges $\mathcal{E}R_X \cap \mathcal{E}R'$. For each boundary edge $e \in \mathcal{E}R'$ such that $e \notin \mathcal{E}R_X$,



(a) $\nu(X) = 30/91$ (b) $\nu(X'_1) = 30/100$ (c) $\nu(X'_2) = 30/100$

Figure 6.4. An example where $\nu(X)$ is decreasing when considering smaller regions. This can only happen when $\lambda > 0$. Here we have $\lambda = 1/2$ and $q = 2$.

let $B_{X'}(e) = B'_{X'}(e) = \theta(v)$, where vertex $v \in \partial R'$ is one endpoint of e. Now,

$$\frac{(1-\lambda)\omega_{c,\theta}}{(1+\lambda)\omega_{c,\theta} + \hat{\omega}_{c,c',\theta}} = \frac{(1-\lambda)\omega'_c}{(1+\lambda)\omega'_c + \hat{\omega}'_{c,c'}} = \mu_{c,c}(X')$$

where ω'_c and $\hat{\omega}'_{c,c'}$ are the weights for the edge-boundary pair X'. The first equality holds since $\omega'_c = \alpha\omega_{c,\theta}$ and $\hat{\omega}'_{c,c'} = \alpha\hat{\omega}_{c,c',\theta}$, where the factor $\alpha = \lambda^{\mathrm{mon}_{\theta,B_X}(E(R_X)\setminus E(R'))}$ is the contribution by edges that do not belong to $E(R')$ and are monochromatic under θ. Finally, note that $X' \in S$. □

The lemma enables us to find upper bounds on $\mu_{c,c'}(X)$. The corresponding lemmas, [GMP05, Lemma 13] and [GJMP06, Lemma 6], are different in that they are stated for the value $\mu(X)$ rather than $\mu_{c,c'}(X)$. However, since $\mu(X)$ is defined as the maximum of $\mu_{c,c'}(X)$ and $\mu_{c',c}(X)$ it follows that also $\mu(X)$ can be upper-bounded as the lemma suggests. We state this fact in the following corollary.

Corollary 6.6. *Suppose that X is an edge-boundary pair. Let R' be any subset of R_X which includes v_X. Let S be the set of edge-boundary pairs X' such that $R_{X'} = R'$, the distinguished edge $e_{X'} = e_X$, and the boundary colouring $B_{X'}$ is identical to B_X, and $B'_{X'}$ is identical to B'_X on the edges $\mathcal{E}R_X \cap \mathcal{E}R'$. Then $\mu(X) \leq \max_{X' \in S} \mu(X')$.*

Lemma 6.5 is stronger than Corollary 6.6 since it guarantees that an upper bound on $\mu_{c,c'}(X)$ can be obtained from a value $\mu_{c,c'}(X')$ where the order of the colours c and c' is preserved. This is a crucial fact we make use of when showing exponential decay of $\Gamma^d(X)$ for the kagome lattice in Chapter 9.

One remark should be mentioned with regards to the values $\nu(X)$ and $\mu(X)$. We want to obtain an upper bound on $\nu(X)$ that is as tight as possible but still feasible to compute. From the proof of Lemma 6.4 it follows that if $\lambda = 0$ then $\nu(X) = \mu(X)$. It is only for positive λ that $\mu(X)$ is strictly greater than $\nu(X)$. This means that Corollary 6.6 above could be stated with $\nu(\cdot)$ instead of $\mu(\cdot)$ for the special case $\lambda = 0$. To see why this would not be correct if $\lambda > 0$, consider the following example with the square lattice, number of colours $q = 2$ and $\lambda = 1/2$. Figure 6.4(a) shows an edge-boundary pair X for

which the region R_X contains two vertices. By explicitly working out the probabilities in π_{B_X} and $\pi_{B'_X}$ for each of the four colourings in $\Omega^+_{R_X}$, it is straightforward calculate that $\nu(X) = 30/91$. Now, if we shrink the region R_X, in accordance with Corollary 6.6, so it only contains vertex v_X, then we have two edge-boundary pairs X'_1 and X'_2 in S for which we have to compute $\nu(\cdot)$. They differ on the colour of the edge that was connecting v_X with the vertex we just removed. The two edge-boundary pairs are shown in Figure 6.4(b) and (c), respectively. By working out the probabilities of colouring v_X with colour 1 and 2, respectively, we find that both $\nu(X'_1) = 30/100$ and $\nu(X'_2) = 30/100$. Thus, each boundary pair X' in S has a value $\nu(X')$ that is *lower* than $\nu(X)$. This is not what is stated by Corollary 6.6 for $\mu(\cdot)$. The conclusion is that $\mu(X)$ as an upper bound on $\nu(X)$ is not necessarily tight. However, as Corollary 6.6 says, an upper bound on $\mu(X)$ can be obtained by considering smaller regions, which is not necessarily possible for $\nu(X)$.

6.3 Exponential decay in general

To see how the value of $\mu(X)$ is linked to the decay of $\Gamma^d(X)$ we give the following lemma, which corresponds to Lemma 18 in [GMP05].

Lemma 6.7. *Suppose there is a constant $0 < \varepsilon < 1$ such that for every edge-boundary pair X, we have*

$$\mu(X) \leq \frac{1-\varepsilon}{\max[1, n_X]}.$$

Then

$$\Gamma^d(X) \leq (1-\varepsilon)^d.$$

Proof. The proof is by induction on d. For the base case, $d = 1$, we have

$$\Gamma^1(X) = \nu(X) \leq \mu(X) \leq (1-\varepsilon).$$

Recall that $\nu(X) \leq \mu(X)$ from Lemma 6.4. For the inductive step we have

$$\begin{aligned}
\Gamma^d(X) &= \sum_{\substack{c,c' \in Q,\\ c \neq c'}} p_X^{\min}(c,c') \sum_{i=1}^{n_X} \Gamma^{d-1}(X_i(c,c')) \leq \sum_{\substack{c,c' \in Q,\\ c \neq c'}} p_X^{\min}(c,c') \sum_{i=1}^{n_X} (1-\varepsilon)^{d-1} \\
&= \nu(X) \cdot n_X \cdot (1-\varepsilon)^{d-1} \leq \mu(X) \cdot n_X \cdot (1-\varepsilon)^{d-1} \\
&\leq \frac{1-\varepsilon}{n_X} \cdot n_X \cdot (1-\varepsilon)^{d-1} = (1-\varepsilon)^d,
\end{aligned}$$

where the first equality is from the definition of $\Gamma^d(X)$ in Equation (6.2) on page 67, and the first inequality uses the induction hypothesis. □

Lemma 6.8 below gives an upper bound on $\mu(X)$ for any edge-boundary pair

$q = 5$ $q = 9$

Figure 6.5. Two examples of boundary colourings that maximise $\mu(X)$ when the region R_X consists of only a single vertex.

X. The upper bound clearly depends on the maximum degree Δ of the graph, the number of colours q and λ. We will use the lemma in order to prove Theorem 3.1, for which we give a full proof shortly.

Let X be an edge-boundary pair. From Corollary 6.6 we have that an upper bound on $\mu(X)$ can be obtained by maximising $\mu(X')$ over all edge-boundary pairs X' for which the region $R_{X'}$ consists only of vertex $v_{X'}$. Suppose X is an edge-boundary pair for which R_X consists only of vertex v_X. In order to find the boundary colourings B_X and B'_X that maximises $\mu(X)$, it turns out that choosing v_X to be a vertex with maximum degree Δ, and spreading colours evenly over the boundary edges $\mathcal{E}R_X$ will maximise $\mu(X)$. More precisely, if the colour of edge e_X is 1 and 2 in B_X and B'_X, respectively, then spreading the colours $3, \ldots, q$ evenly over the remaining $\Delta - 1$ edges will maximise $\mu(X)$. That is, each colour in $\{3, \ldots, q\}$ is first assigned to $a = \lfloor (\Delta - 1)/(q - 2) \rfloor$ edges. Now there are $r = (\Delta - 1) \bmod (q - 2) < q - 2$ edges left to be assigned a colour. These r edges are given distinct colours in $\{3, \ldots, q\}$. See Figure 6.5 for two concrete examples. We have the following lemma.

Lemma 6.8. *Consider the anti-ferromagnetic Potts model on an infinite graph $\mathcal{G}$ with maximum degree Δ and parameters q and $\lambda \in [0, 1]$. Let $a = \lfloor (\Delta - 1)/(q - 2) \rfloor$, $r = (\Delta - 1) \bmod (q - 2)$ (so $a(q - 2) + r = \Delta - 1$) and*

$$\kappa = \frac{1 - \lambda}{1 + \lambda + r\lambda^{a+1} + (q - 2 - r)\lambda^a}.$$

Then $\mu(X) \leq \kappa$ for every edge-boundary pair X.

Proof. From Corollary 6.6 it follows that $\mu(X)$ is maximised if X is an edge-boundary pair for which the region R_X consists of only vertex v_X. Let X be such an edge-boundary pair. We want to choose two colourings B_X and B'_X that maximise $\mu(X)$. Without loss of generality, suppose that $B_X(e_X) = 1$ and $B'_X(e_X) = 2$, and suppose that the weight $\omega_1 \geq \omega_2$. Let $E_X = \mathcal{E}R_X \backslash \{e_X\}$ be the set of edges that are incident to v_X, excluding edge e_X. Note that $|E_X| \leq \Delta - 1$. For $i \in Q$, let n_i be the number of edges in E_X that are assigned colour i by B_X. Note that $\omega_i = \lambda^{n_i}$, so the constraint $\omega_1 \geq \omega_2$ just says $n_2 \geq n_1$.

Now we wish to choose B_X in order to maximise $\mu(X)$. Since $\omega_1 \geq \omega_2$, we have

$$\mu(X) = \frac{(1-\lambda)\omega_1}{(1+\lambda)\omega_1 + \hat{\omega}_{1,2}} = \frac{1}{\frac{1+\lambda}{1-\lambda} + \frac{\hat{\omega}_{1,2}}{(1-\lambda)\omega_1}}.$$

So equivalently to maximise $\mu(X)$, we wish to minimise

$$Z = \frac{\hat{\omega}_{1,2}}{(1-\lambda)\omega_1}.$$

First note that $n_1 = 0$, and hence $\omega_1 = 1$, since Z can be reduced by recolouring edges that are coloured 1 with 2. This does not affect $\hat{\omega}_{1,2}$.

Next we want to set $n_2, \ldots, n_q$ in order to minimise

$$\hat{\omega}_{1,2} = \lambda^{n_3} + \cdots + \lambda^{n_q},$$

where $n_3 + \cdots + n_q \leq \Delta - 1$. Since $\lambda \leq 1$, we want to take $n_3 + \cdots + n_q = \Delta - 1$. That is, in order to find the maximum value of $\mu(X)$, we let X be an edge-boundary pair for which $|E_X| = \Delta - 1$.

It remains to distribute $\Delta - 1$ over $n_3, \ldots, n_q$. In order to minimise $\hat{\omega}_{1,2}$ we want to spread the colours evenly over the edges in E_X. That is, there is an optimal solution in which all n_j and n_k are within 1 of each other. To see this, consider a solution with $n_j \geq n_k + 2$. The boundary obtained by reassigning the colour of one of the edges coloured j with k has a $\hat{\omega}_{1,2}$-value that is not bigger since the new $\hat{\omega}_{1,2}$-value minus the old one is

$$(\lambda^{n_j - 1} + \lambda^{n_k + 1}) - (\lambda^{n_j} + \lambda^{n_k}) = (1-\lambda)(\lambda^{n_j - 1} - \lambda^{n_k}) \leq 0.$$

So the optimum value of $\hat{\omega}_{1,2}$ is $r\lambda^{a+1} + (q-2-r)\lambda^a$, which together with $\omega_1 = 1$ proves the lemma. □

We now have all the pieces we need in order to give a complete proof of Theorem 3.1. We repeat the theorem and give the proof.

Theorem (3.1). *Consider the anti-ferromagnetic Potts model on an infinite and uniformly sub-exponential graph $\mathcal{G}$ with maximum degree Δ and parameters q and $\lambda \in [0,1]$. Let $a = \lfloor (\Delta - 1)/(q-2) \rfloor$, $r = (\Delta - 1) \bmod (q-2)$ (so $a(q-2) + r = \Delta - 1$) and*

$$\kappa = \frac{1-\lambda}{1 + \lambda + r\lambda^{a+1} + (q-2-r)\lambda^a}.$$

The system has strong spatial mixing if $\kappa < 1/(\Delta - 1)$. Furthermore, if $\kappa < 1/(\Delta - 1)$, and $q \geq \Delta + 2$ or $\lambda > 0$, then the Glauber dynamics is rapidly mixing for any region R and any q_0-colouring $\mathcal{B}$ of ∂R. The mixing time $\tau(\delta) \in O(n^2 + n \log \frac{1}{\delta})$, where n is the

number of vertices in R.

Note, if $q \geq \Delta + 1$ and $q > (2\Delta - 1)(1 - \lambda)$ then $\kappa < 1/(\Delta - 1)$.

Proof. Suppose that there is a Δ, q and λ such that $\kappa < 1/(\Delta - 1)$. Since κ is a constant, there must be an $\varepsilon > 0$ such that

$$\kappa \leq \frac{1 - \varepsilon}{\Delta - 1}.$$

From Lemma 6.8 we have

$$\mu(X) \leq \kappa \leq \frac{1 - \varepsilon}{\Delta - 1}$$

for every edge-boundary pair X. We apply Lemma 6.7, which gives that $\Gamma^d(X) \leq (1 - \varepsilon)^d$ for every edge-boundary pair X. Using Lemma 6.3 gives us strong spatial mixing and the mixing time of the Glauber dynamics.

For the last statement of the theorem, suppose that $q > \Delta + 1$. Then $a = 0$ and $r = \Delta - 1$, which implies that

$$\kappa = \frac{1 - \lambda}{1 + \lambda + (\Delta - 1)\lambda + (q - 1 - \Delta)}. \tag{6.3}$$

If $q = \Delta + 1$ then $a = 1$ and $r = 0$, which implies that

$$\kappa = \frac{1 - \lambda}{1 + \lambda + (q - 2)\lambda}. \tag{6.4}$$

Now, if $q > (2\Delta - 1)(1 - \lambda)$ it is straightforward to verify that the right hand side of Equation (6.3) and (6.4) are both strictly less than $1/(\Delta - 1)$. □

From Theorem 3.1 it follows that the system has strong spatial mixing if $q > 2\Delta - 1$ and $\lambda = 0$. As mentioned in Section 3.1, Goldberg, Martin and Paterson [GMP05] have provided a better result than this for $\lambda = 0$. Their result is stated in Theorem 3.8. We will not repeat the proof of their theorem here, but we will briefly explain the idea behind it. The idea is very similar to the proof of Lemma 6.8 but is more involved. We consider edge-boundary pairs X with small regions R_X and construct colourings B_X and B'_X that would maximise $\mu(X)$. In the proof of Lemma 6.8 we considered regions R_X that contained the single vertex v_X. In order to prove Theorem 3.8, Goldberg, Martin and Paterson considered regions R_X that contain v_X and neighbours of v_X. They showed that if the vertex v_X has $n_X \geq 1$ neighbours in R_X then $\mu(X) < 1/n_X$. By applying Lemma 6.7, exponential decay of $\Gamma^d(X)$ follows.

Let X be an edge-boundary pair for which R_X contains vertex v_X and $n_X \geq 1$ neighbours $v_1, \ldots, v_{n_X}$. Note that Theorem 3.8 holds only for triangle-free graphs, so there are no edges connecting any pair of vertices in R_X, except for the edges between v_X and its neighbours. The reason why the theorem is restricted to triangle-free graphs is because otherwise the increased interaction between neighbouring vertices in R_X

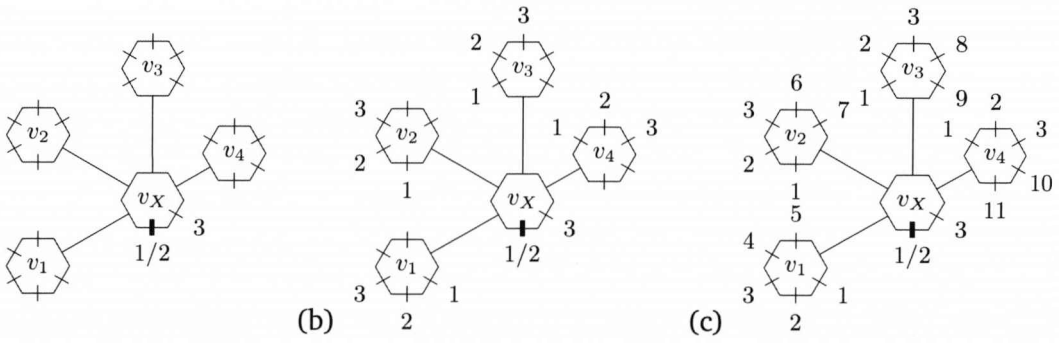


Figure 6.6. An example of a boundary colouring (c) that maximises $\mu(X)$ when $q = 13$, $\lambda = 0$ and the region R_X consists of vertex v_X and four of its neighbours. Note that we draw the neighbours of v_X slightly apart to give room for the labels. Also, the graph is supposed to be triangle-free.

would make the system much harder to analyse. The value $\mu(X)$ will be maximised if all vertices in R_X are incident to as many boundary edges as possible. That is, each vertex in R_X has degree Δ. The following procedure will assign colours to the boundary edges such that $\mu(X)$ is maximised. We refer to Figure 6.6 for a concrete example, where $\Delta = 6$, $q = 13$ and v_X has $n_X = 4$ neighbours in R_X.

1. Without loss of generality, suppose $B_X(e_X) = 1$, $B'_X(e_X) = 2$. See Figure 6.6(a).

2. The edges in $\mathcal{E}R_X \backslash \{e_X\}$ that are incident to v_X are assigned colour $3, \ldots, (\Delta - n_X + 1)$, respectively. See Figure 6.6(a), in which the only edge adjacent to v_X is given the colour 3.

3. For each neighbour v of v_X, edges incident to v are assigned the colours that have been used so far (one edge per colour). That is, $\Delta - n_X + 1$ of the edges incident to v are assigned the colours $1, \ldots, (\Delta - n_X + 1)$, respectively. See Figure 6.6(b).

4. The colours $(\Delta - n_X + 2), \ldots, q$ have still not been used. These colours are now spread evenly over the remaining edges. See Figure 6.6(c), in which the remaining 8 edges are given the colours $4, \ldots, 11$, respectively.

Boundary colourings B_X and B'_X that are constructed according to this procedure will maximise $\mu(X)$. The optimisation is rigorously proved in [GMP05] and shows that $\mu(X) < 1/n_X$ if $q \geq \alpha\Delta - \gamma$, where α is the solution to $\alpha^\alpha = e$ (so $\alpha \approx 1.76322$) and $\gamma = \frac{4\alpha^3 - 6\alpha^2 - 3\alpha + 4}{2(\alpha^2 - 1)} \approx 0.47031$.

6.4 Exponential decay for specific spin systems

In the previous section we stated Lemma 6.7, which says that $\Gamma^d(X)$ is decreasing exponentially if $\mu(X) < 1/n_X$ for all edge-boundary pairs X. This is not a necessary

condition for exponential decay. The quantity $\Gamma^d(X)$ is determined by the weights of the edges in the tree T_X, and if some edges are too "heavy" then this can be compensated by other edges being "lighter". In terms of the value $\mu(X)$, it means that if $\mu(X) > 1/n_X$ then $\Gamma^d(X)$ could still be decreasing exponentially if there are other values $\mu(X')$ that are sufficiently small, where X' is an edge-boundary pair constructed recursively in the tree T_X. In order to capture this fact, we will introduce some notation that might seem rather artificial. As we will see, this notation will be useful in subsequent chapters where we prove exponential decay for various spin systems. These proofs are computer assisted, and the definitions we introduce will provide an easy-to-use framework for the computations. We start with the following definition of an $(\mathcal{A}, \mathcal{F})$-set, which is used to describe how edge-boundary pairs are related in the recursive-coupling tree T_X. The idea behind the notation will be explained after the definition.

Definition 6.9 ($(\mathcal{A}, \mathcal{F})$-set). *Let $\mathcal{A}$ and $\mathcal{F}$ be two collections of sets of edge-boundary pairs. A subset*

$$S \subseteq \mathcal{A} \times \mathcal{F} \times 2^{(\mathcal{A} \cup \{\emptyset\})^{\Delta - 1}}$$

is an $(\mathcal{A}, \mathcal{F})$-set, where Δ is the maximum degree of the graph $\mathcal{G}$.

Let $\mathcal{A}$ be a collection of sets of edge-boundary pairs, which do not need to be disjoint, but the union of them have to contain all edge-boundary pairs. Similarly, let $\mathcal{F}$ be a collection of sets of edge-boundary pairs, not necessarily disjoint, but the union of them contains all edge-boundary pairs. An edge-boundary pair X is associated with an element $(A, F, \mathcal{H}) \in S$. The first component A is a set in $\mathcal{A}$ to which X belongs. Note that X can belong to several of the sets in $\mathcal{A}$. Similarly, the second component F is a set in $\mathcal{F}$ to which X also belongs. The roles of A and F differ; A will be used to classify X, and F will be used to obtain an upper bound on $\mu(X)$. The third component, $\mathcal{H}$, is a set of $(\Delta - 1)$-tuples $(A_1, \ldots, A_{\Delta-1})$, such that A_i is a set in $\mathcal{A}$ or is the empty set $\emptyset$. In the construction of the tree T_X, we derive edge-boundary pairs $X_i(c, c')$ from X, where $i = 1, \ldots, n_X$ and $c, c' \in Q$ are two distinct colours such that $p_X^{\min}(c, c') > 0$. For a fixed pair $(c, c') \in Q \times Q$ of colours, such that $c \neq c'$ and $p_X^{\min}(c, c') > 0$, there is a tuple $H = (A_1, \ldots, A_{\Delta-1}) \in \mathcal{H}$ that contains sets in $\mathcal{A}$ to which $X_1(c, c'), \ldots, X_{n_X}(c, c')$ belong. The ordering of $A_1, \ldots, A_{\Delta-1}$ is not important, but each of the $\Delta - 1$ sets in H can only be associated with one of the edge-boundary pairs $X_i(c, c')$. That is, if n of the n_X edge-boundary pairs $X_i(c, c')$ belong to $A \in \mathcal{A}$, and if we want to associate them with A, then n distinct elements in $(A_1, \ldots, A_{\Delta-1})$ are equal to the set A.

The idea of the set $\mathcal{H}$ is that each tuple $(A_1, \ldots, A_{\Delta-1})$ in $\mathcal{H}$ is associated with a choice c, c' of colours and each A_j is a set in $\mathcal{A}$ to which an edge-boundary pair $X_i(c, c')$ belongs. The colours c, c' are distinct colours in Q with $p_X^{\min}(c, c') > 0$. The following definition of a *valid* $(\mathcal{A}, \mathcal{F})$-set formally summarises the idea.

Definition 6.10 (Valid $(\mathcal{A}, \mathcal{F})$-set). *An $(\mathcal{A}, \mathcal{F})$-set S is* valid *if*

(1) $\mathcal{A}$ and $\mathcal{F}$ are two collections of sets of edge-boundary pairs, such that any edge-boundary pair X belongs to at least one set in $\mathcal{A}$ and to at least one set in $\mathcal{F}$, and

(2) for every set $A \in \mathcal{A}$ and every edge-boundary pair $X \in A$, there is an element $(A, F, \mathcal{H})$ in S such that $X \in F$, and for all distinct colours $c, c' \in Q$ for which $p_X^{\min}(c, c') > 0$, there is a tuple $(A_1, \ldots, A_{\Delta-1}) \in \mathcal{H}$ and an injective function $g : \{1, \ldots, n_X\} \to \{1, \ldots, \Delta - 1\}$, such that $X_i(c, c') \in A_{g(i)}$, where $i \in \{1, \ldots, n_X\}$ and $X_i(c, c')$ is the edge-boundary pair constructed recursively in the tree T_X.

The notion of a *good* $(\mathcal{A}, \mathcal{F})$-set S below is linking S to some values μ_F, where F is a set in $\mathcal{F}$. The value μ_F is an upper bound on $\mu(X)$ over all $X \in F$. By using different upper bounds on $\mu(X)$ for different edge-boundary pairs X, we can afford some values being "too large", if they are compensated by other values being "sufficiently small". The word "good" here is referring to the fact that the values μ_F enables us to show exponential decay of $\Gamma^d(X)$.

Definition 6.11 (Good $(\mathcal{A}, \mathcal{F})$-set). *Let S be a valid $(\mathcal{A}, \mathcal{F})$-set, and for each set $F \in \mathcal{F}$, let μ_F be a value in $[0, 1]$. The set S is* good with respect to a constant $\varepsilon \in (0, 1)$ and the values μ_F *if the following is true:*

(1) For every $F \in \mathcal{F}$,

$$\max_{X \in F} \mu(X) \leq \mu_F.$$

(2) For each set $A \in \mathcal{A}$ there is a constant $\alpha_A \geq 1/(1 - \varepsilon)$ such that for every element $(A, F, \mathcal{H}) \in S$, and every tuple $(A_1, \ldots, A_{\Delta-1}) \in \mathcal{H}$,

$$\mu_F(\alpha_{A_1} + \cdots + \alpha_{A_{\Delta-1}}) \leq \alpha_A(1 - \varepsilon),$$

where $\alpha_\emptyset = 0$.

The next lemma relates good $(\mathcal{A}, \mathcal{F})$-sets to exponential decay of $\Gamma^d(X)$.

Lemma 6.12. *Let S be a valid $(\mathcal{A}, \mathcal{F})$-set that is good with respect to a constant $\varepsilon \in (0, 1)$ and some values μ_F, where $F \in \mathcal{F}$. Then there is a constant $\alpha \geq 0$ such that $\Gamma^d(X) \leq \alpha(1 - \varepsilon)^d$ for all edge-boundary pairs X.*

Proof. For any set A of edge-boundary pairs, let $\Gamma^d(A)$ denote the maximum of $\Gamma^d(X)$ over all $X \in A$. Let $\Gamma^d(\emptyset) = 0$ for all $d \geq 1$. In order to show that $\Gamma^d(X) \leq \alpha(1-\varepsilon)^d$ for every edge-boundary pair X, we will show that $\Gamma^d(A) \leq \alpha_A(1 - \varepsilon)^d$ for every $A \in \mathcal{A}$, and take α as the maximum of α_A over all $A \in \mathcal{A}$. Note that any edge-boundary pair X belongs to at least one of the sets in $\mathcal{A}$.

The proof is by induction on d. We start with the base case $d = 1$. Consider $A \in \mathcal{A}$. We wish to show $\Gamma^1(A) \leq \alpha_A(1-\varepsilon)$. Consider an edge-boundary pair $X \in A$. Since $\alpha_A \geq 1/(1-\varepsilon)$, we have

$$\Gamma^1(X) = \nu(X) \leq \mu(X) \leq 1 \leq \alpha_A(1-\varepsilon),$$

where the first inequality is from Lemma 6.4. Now consider the inductive step for $d > 1$. Consider $X \in A$ for $A \in \mathcal{A}$. We wish to show $\Gamma^d(A) \leq \alpha_A(1-\varepsilon)^d$. Now by Equation (6.2) on page 67 we have

$$\Gamma^d(X) = \sum_{\substack{c,c' \in Q,\\ c \neq c'}} p_X^{\min}(c,c') \sum_{i=1}^{n_X} \Gamma^{d-1}(X_i(c,c')). \tag{6.5}$$

Let $(A, F, \mathcal{H})$ be a set in S such that condition (2) in Defintion 6.10 holds. For every pair of distinct colours $c, c' \in Q$ for which $p_X^{\min}(c,c') > 0$, we know that there is a tuple $(A_1, \ldots, A_{\Delta-1}) \in \mathcal{H}$ and an injective function $g : \{1, \ldots, n_X\} \to \{1, \ldots, \Delta - 1\}$, such that $X_i(c,c') \in A_{g(i)}$, where $i = 1, \ldots, n_X$ and $X_i(c,c')$ is the edge-boundary pair constructed recursively in the tree T_X. We have,

$$\sum_{i=1}^{n_X} \Gamma^{d-1}(X_i(c,c')) \leq \sum_{i=1}^{n_X} \Gamma^{d-1}(A_{g(i)}) \leq \sum_{i=1}^{\Delta-1} \Gamma^{d-1}(A_i) \leq \sum_{i=1}^{\Delta-1} \alpha_{A_i}(1-\varepsilon)^{d-1}, \tag{6.6}$$

where the last inequality uses the induction hypothesis. In the second last inequality we sum over all elements in the tuple $(A_1, \ldots, A_{\Delta-1})$ and not only over those n_X distinct elements mapped to by the injective function g. Since $X \in F$ we have $\mu(X) \leq \mu_F$. Suppose that the tuple $(A_1, \ldots, A_{\Delta-1}) \in \mathcal{H}$ is maximising the right-hand side of Equation (6.6) over all tuples in $\mathcal{H}$. Then using Equation (6.6) with Equation (6.5) gives

$$\begin{aligned}
\Gamma^d(X) &\leq \sum_{\substack{c,c' \in Q,\\ c \neq c'}} p_X^{\min}(c,c') \sum_{i=1}^{\Delta-1} \alpha_{A_i}(1-\varepsilon)^{d-1} \\
&= \sum_{\substack{c,c' \in Q,\\ c \neq c'}} p_X^{\min}(c,c') \cdot (\alpha_{A_1} + \cdots + \alpha_{A_{\Delta-1}})(1-\varepsilon)^{d-1} \\
&= \nu(X)(\alpha_{A_1} + \cdots + \alpha_{A_{\Delta-1}})(1-\varepsilon)^{d-1} \leq \mu_F(\alpha_{A_1} + \cdots + \alpha_{A_{\Delta-1}})(1-\varepsilon)^{d-1} \\
&\leq \alpha_A(1-\varepsilon)^d,
\end{aligned}$$

where the last inequality follows from the fact that S is good with respect to ε and the values μ_F, and hence $\mu_F(\alpha_{A_1} + \cdots + \alpha_{A_{\Delta-1}}) \leq \alpha_A(1-\varepsilon)$. □

The problem of showing exponential decay of $\Gamma^d(X)$ can be reduced to the problem of defining two collections $\mathcal{A}$ and $\mathcal{F}$ of sets of edge-boundary pairs, deriving values μ_F for the sets $F \in \mathcal{F}$, and constructing a valid $(\mathcal{A}, \mathcal{F})$-set that is good with respect to the values μ_F. This process can be done in a systematic way. With the help of a computer it is possible to work with rather large number of sets in $\mathcal{A}$ and $\mathcal{F}$, and consider $(\mathcal{A}, \mathcal{F})$-sets that contain thousands of elements. Also, with computer assistance, $\mu(X)$ can be calculated for edge-boundary pairs X for which the region R_X contains more vertices than only v_X and its neighbours.

In practice, the constants α_A in Definition 6.11 are obtained by solving a linear program, for which the constraints are given by the inequalities that have to be satisfied by the $(\mathcal{A}, \mathcal{F})$-set in order for it to be good with respect to the values μ_F. That is, for a fixed choice of $\varepsilon > 0$, the inequalities are $\alpha_A \geq 1/(1-\varepsilon)$, $A \in \mathcal{A}$, and for every element $(A, F, \mathcal{H}) \in S$ and every tuple $(A_1, \ldots, A_{\Delta-1}) \in \mathcal{H}$, we have the inequality $\mu_F(\alpha_{A_1} + \cdots + \alpha_{A_{\Delta-1}}) \leq \alpha_A(1-\varepsilon)$. Note that solving a linear program is not a goal in itself. We are only interested in a solution to the set of inequalities. Therefore we can optimise any objective function, subject to the set of inequalities. As we will see in subsequent chapters, the number of inequalities we want to satisfy is rather large. Using a large-scale linear program solver on a computer will provide a useful tool when searching for a solution to the set of inequalities.

6.5 Extended regions

When constructing $(\mathcal{A}, \mathcal{F})$-sets it will be convenient to introduce the notion of an *extended region* $\mathcal{R}$, which is a region with the following additional information:

- Every vertex in $\mathcal{R}$ is labelled either "in" or "out".
- One of the boundary edges of $\mathcal{R}$ is referred to as the designated edge.

When illustrating extended regions $\mathcal{R}$ in the figures, we let non-shaded faces represent vertices in $\mathcal{R}$ that are labelled "in", and we let shaded faces represent vertices in $\mathcal{R}$ that are labelled "out". We mark the designated boundary edge with a short and thick line segment. Figure 6.7(a) illustrates an extended region $\mathcal{R}$ of the square lattice, consisting of three vertices that are labelled "in", and one vertex that is labelled "out". A region R and an extended region $\mathcal{R}$ are *matching* with respect to an edge $e \in E(R)$ if there is a way of overlapping R with $\mathcal{R}$ such that the designated edge of $\mathcal{R}$ coincides with the edge e, and every vertex that is labelled "in" in $\mathcal{R}$ coincides with a vertex that is in R, and every vertex that is labelled "out" in $\mathcal{R}$ coincides with a vertex that is not in R. For example, the extended region in Figure 6.7(a) is matching the region in Figure 6.7(b) with respect to the edge that is marked in the figure. The overlap is illustrated in

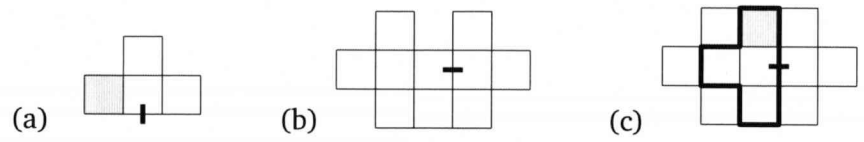


Figure 6.7. (a) An example of how an extended region is illustrated. A non-shaded vertex is "in", and a shaded vertex is "out". In (c) we see that the extended region in (a) matches the region in (b) with respect to the edge that is marked in (b).

Figure 6.7(c). Note that the overlapping takes place under any rotation or reflection of the regions.

We will use extended regions to define sets of edge-boundary pairs. For instance, we could define a set A to be the set of all edge-boundary pairs X such that the region R_X is matching the extended region in Figure 6.7(a) with respect to edge e_X. Then A would contain all edge-boundary pairs X on the square lattice, for which v_X has exactly two neighbours v_1 and v_2 in R_X, such that v_1 and v_2 are both neighbours to a vertex v_3 that is not the same vertex as v_X.

Chapter 7

The square lattice

In this chapter we prove Theorem 3.2. We define a collection $\mathcal{A}$ of seven sets $A_1, \ldots, A_7$ of edge-boundary pairs, and a collection $\mathcal{F}$ of eight sets $F_1, \ldots, F_8$ of edge-boundary pairs. We define a valid $(\mathcal{A}, \mathcal{F})$-set S which contains no more than 47 elements. It is straightforward to verify by hand that the set S is valid. In order to show exponential decay of $\Gamma^d(X)$ for every edge-boundary pair X, we compute eight values $\mu_{F_1}, \ldots, \mu_{F_8}$ and show that S is good with respect to these values. The values clearly depend on the number of colours q. We consider $q = 3, \ldots, 6$. The values $\mu_{F_1}, \ldots, \mu_{F_8}$, which are upper bounds on $\mu(X)$ for all $X \in F_1, \ldots, F_8$, respectively, are derived with the help of a computer, which makes the proof of Theorem 3.2 computer assisted.

The organisation of the chapter is as follows. In Section 7.1 we define the collection $\mathcal{F}$ of sets of edge-boundary pairs and present the values μ_F for each $F \in \mathcal{F}$. In Section 7.2 we define the collection $\mathcal{A}$ of edge-boundary pairs and present a valid $(\mathcal{A}, \mathcal{F})$-set. In Section 7.3 we show that this $(\mathcal{A}, \mathcal{F})$-set is good with respect to the values μ_F, and we give the proof of Theorem 3.2. Finally, in Section 7.4 we discuss some of the issues we have to deal with when implementing the computational part of the proof.

7.1 Upper-bounding $\mu(X)$

For a fixed number of colours q, we note that the quantity $\mu(X)$ is a rational function in λ. For example, let $q = 5$ and consider the edge-boundary pair X illustrated in Figure 7.1. For this edge-boundary pair we have the weights $\omega_1 = \lambda$, $\omega_2 = \lambda^2$ and $\omega_3 = \omega_4 = \omega_5 = 1$. Thus,

$$\mu(X) = \max[\mu_{1,2}, \mu_{2,1}] = \frac{(1-\lambda)\omega_1}{(1+\lambda)\omega_1 + \hat{\omega}_{1,2}} = \frac{(1-\lambda)\lambda}{(1+\lambda)\lambda + 3} = \frac{\lambda - \lambda^2}{\lambda + \lambda^2 + 3}. \qquad (7.1)$$

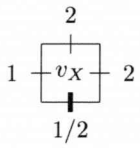


Figure 7.1. An edge-boundary pair X. The value $\mu(X) = (\lambda - \lambda^2)/(\lambda + \lambda^2 + 3)$ when $q = 5$.

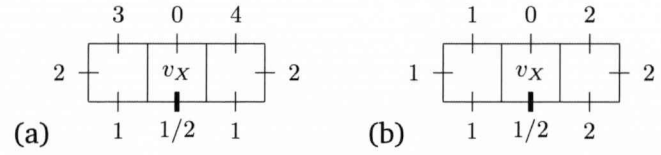


Figure 7.2. Two edge-boundary pairs X (a) and X' (b) such that $\mu(X) > \mu(X')$ when $\lambda = 0$ and $q = 4$, but $\mu(X) < \mu(X')$ when $\lambda = 1/2$ and $q = 4$.

In general, $\mu(X)$ is not monotonically decreasing in λ, which is confirmed by the value of $\mu(X)$ in Equation (7.1). Here $\mu(X) = 0$ for both $\lambda = 0$ and $\lambda = 1$, but $\mu(X) = 1/15$ for $\lambda = 1/2$. Hence we cannot assume that $\lambda = \lambda_0$ is maximising $\mu(X)$ on the interval $\lambda \in [\lambda_0, 1]$, where $\lambda_0 \in [0, 1]$ is a constant. Nevertheless, for a fixed region R and boundary edge $e \in \mathcal{E}R$, it seems like $\max_X \mu(X)$ is monotonically decreasing in λ, where the maximum is taken over all edge-boundary pairs X such that $R_X = R$ and $e_X = e$.

Suppose X is an edge-boundary pair. In order to obtain an upper bound on $\mu(X)$ for $\lambda \in [\lambda_{\mathsf{min}}, \lambda_{\mathsf{max}}]$, we divide the interval $[\lambda_{\mathsf{min}}, \lambda_{\mathsf{max}}]$ into $m \geq 1$ intervals. Let $\lambda_0 = \lambda_{\mathsf{min}}$ and $\lambda_m = \lambda_{\mathsf{max}}$. We consider the m intervals $[\lambda_0, \lambda_1]$, $[\lambda_1, \lambda_2]$, $[\lambda_2, \lambda_3], \ldots,$ $[\lambda_{m-1}, \lambda_m]$. For every $i \in \{0, \ldots, m-1\}$, let μ_i be an upper bound on $\mu(X)$ for $\lambda \in [\lambda_i, \lambda_{i+1}]$. The upper bound μ_i is the value of $\mu(X)$ computed by taking $\lambda = \lambda_i$ for negative terms in the numerator of $\mu(X)$ and $\lambda = \lambda_{i+1}$ for positive terms in the numerator. We take $\lambda = \lambda_i$ for the terms in the denominator since they are all positive. An upper bound on $\mu(X)$ for $\lambda \in [\lambda_{\mathsf{min}}, \lambda_{\mathsf{max}}]$ is then obtained by maximising μ_i over $i \in \{0, \ldots, m-1\}$.

We make the following remark. For two edge-boundary pairs X and X', such that $R_X = R_{X'}$ and $e_X = e_{X'}$, the quantity $\mu(X)$ could be strictly less than $\mu(X')$ for some λ, and strictly greater than $\mu(X')$ for some other λ. Thus, there is not necessarily some "worst" edge-boundary pair X such that $\mu(X) \geq \mu(X')$ for all $\lambda \in [0, 1]$ and all edge-boundary pairs X', where X and X' only differ on the colours of the boundary edges. The two edge-boundary pairs in Figure 7.2 confirm this fact. Suppose $q = 4$ and let X be the edge-boundary pair in Figure 7.2(a). The weights for X are $\omega_1 = \omega_2 = (\lambda^2 + 2\lambda + 1)^2$ and $\omega_3 = \omega_4 = (\lambda^2 + 2\lambda + 1)(4\lambda)$. Let X' be the edge-boundary pair in Figure 7.2(b), which has the weights $\omega_1 = \omega_2 = (\lambda^4 + 3)(\lambda^3 + \lambda + 2)$ and $\omega_3 = \omega_4 = (\lambda^3 + \lambda + 2)^2$. Now, if $\lambda = 0$ then we have $\mu(X) = 1$ and $\mu(X') = 6/14$, so

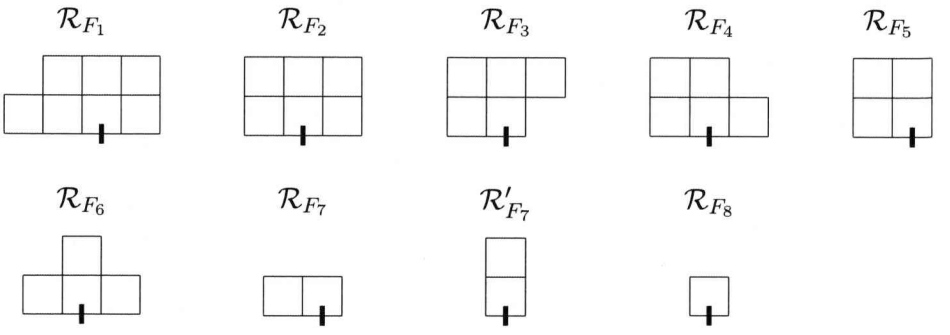


Figure 7.3. Extended regions we use when defining the collection $\mathcal{F}$ of sets of edge-boundary pairs. All vertices are labelled "in".

$\mu(X) > \mu(X')$. If $\lambda = 1/2$ then $\mu(X) = 9/59$ and $\mu(X') = 7/45$, so $\mu(X) < \mu(X')$.

We define a collection $\mathcal{F}$ of eight sets $F_1, \ldots, F_8$ of edge-boundary pairs X. For $i = 1, \ldots, 8$, the set F_i contains all edge-boundary pairs X such that the region R_X is matching the extended region $\mathcal{R}_{F_i}$ in Figure 7.3 with respect to edge e_X. Note that F_7 is defined to contain all edge-boundary pairs X such that the region R_X is matching either the extended region $\mathcal{R}_{F_7}$ or $\mathcal{R}'_{F_7}$ with respect to e_X. That is, F_7 contains all edge-boundary pairs X such that v_X has at least one neighbour in R_X.

Fix a number of colours q and an interval $[\lambda_{\min}, \lambda_{\max}]$. For $i = 1, \ldots, 8$, we want to compute a value μ_{F_i} such that for any $\lambda \in [\lambda_{\min}, \lambda_{\max}]$ and any edge-boundary pair $X \in F_i$, $\mu(X) \leq \mu_{F_i}$. In order to do so, we use Corollary 6.6 and consider every edge-boundary pair $X \in F_i$ for which the region R_X is identical to the extended region $\mathcal{R}_{F_i}$. Each such edge-boundary pair consists of a pair (B_X, B'_X) of boundary colourings, from which we derive $\mu(X)$ as a rational function in λ. As described in the previous section, an upper bound $\mu_{\text{up}}(X)$ on $\mu(X)$ for $\lambda \in [\lambda_{\min}, \lambda_{\max}]$ is obtained by using a suitable partition of the interval $[\lambda_{\min}, \lambda_{\max}]$. The value μ_{F_i} is obtained by maximising $\mu_{\text{up}}(X)$ over all pairs (B_X, B'_X) of boundary colourings. Doing this by hand for all eight regions is not feasible. Instead we let a computer calculate μ_{F_i} by looping through boundary colourings of the regions (see Section 7.4). The upper bounds on μ_{F_i} that we obtain are used in the following lemmas.

Lemma 7.1. *Suppose $q = 6$ and $\lambda \in [0, 1]$. For $i \in \{1, \ldots, 8\}$ and any edge-boundary pair $X \in F_i$, we have $\mu(X) \leq \mu_{F_i}$, where $\mu_{F_1} = 20559/59000$, $\mu_{F_2} = 179501/501000$, $\mu_{F_3} = 4951/13500$, $\mu_{F_4} = 37601/101000$, $\mu_{F_5} = 49129/129000$, $\mu_{F_6} = 27071/71000$, $\mu_{F_7} = 3007/7000$ and $\mu_{F_8} = 1$.*

Lemma 7.2. *Suppose $q = 5$ and $\lambda \in [0.127, 0.286]$. For $i \in \{1, \ldots, 8\}$ and any edge-boundary pair $X \in F_i$, we have $\mu(X) \leq \mu_{F_i}$, where $\mu_{F_1} = 351/1000$, $\mu_{F_2} = 361/1000$, $\mu_{F_3} = 381/1000$, $\mu_{F_4} = \mu_{F_5} = \mu_{F_6} = 401/1000$, $\mu_{F_7} = 501/1000$ and $\mu_{F_8} = 1$.*

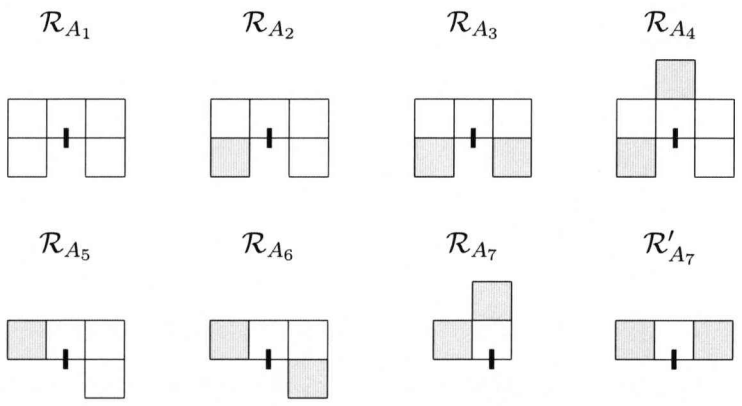


Figure 7.4. Extended regions we use when defining the collection $\mathcal{A}$ of sets of edge-boundary pairs.

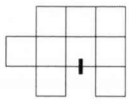
Lemma 7.3. *Suppose $q = 4$ and $\lambda \in [0.262, 0.373]$. For $i \in \{1, \ldots, 8\}$ and any edge-boundary pair $X \in F_i$, we have $\mu(X) \leq \mu_{F_i}$, where $\mu_{F_1} = 351/1000$, $\mu_{F_2} = \mu_{F_3} = \mu_{F_4} = \mu_{F_6} = 381/1000$, $\mu_{F_5} = 401/1000$, $\mu_{F_7} = 501/1000$ and $\mu_{F_8} = 1$.*

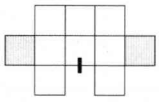
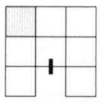
Lemma 7.4. *Suppose $q = 3$ and $\lambda \in [0.393, 0.474]$. For $i \in \{1, \ldots, 8\}$ and any edge-boundary pair $X \in F_i$, we have $\mu(X) \leq \mu_{F_i}$, where $\mu_{F_1} = 1751/5000$, $\mu_{F_2} = \mu_{F_4} = 361/1000$, $\mu_{F_3} = 77/200$, $\mu_{F_5} = 391/1000$, $\mu_{F_6} = 371/1000$, $\mu_{F_7} = 501/1000$ and $\mu_{F_8} = 1$.*

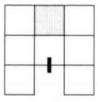
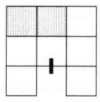
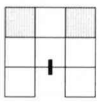
7.2 Defining a valid $(\mathcal{A}, \mathcal{F})$-set

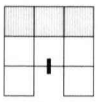
We define a collection $\mathcal{A}$ of seven sets $A_1, \ldots, A_7$ of edge-boundary pairs X. For $i \in \{1, \ldots, 7\}$, the set A_i contains all edge-boundary pairs X such that the region R_X is matching the extended region $\mathcal{R}_{A_i}$ in Figure 6.7 with respect to edge e_X. Note that A_7 is defined to contain all edge-boundary pairs X such that the region R_X is matching either the extended region $\mathcal{R}_{A_7}$ or $\mathcal{R}'_{A_7}$ with respect to e_X.

We construct a valid $(\mathcal{A}, \mathcal{F})$-set S. The elements in S are listed in Figure 7.5–7.6. It is straightforward, although rather tedious, to verify that this set is valid. For instance, let X be any edge-boundary pair in A_1. Then R_X matches, with respect to edge e_X, exactly one of the six extended regions listed under A_1 in Figure 7.5. Let $\mathcal{R}$ be the extended region in the top right corner of Figure 7.5. Suppose that R_X matches $\mathcal{R}$. We see that X belongs to F_4. For any pair of distinct colours $c, c' \in Q$, the edge-boundary pairs $X_1(c, c')$, $X_2(c, c')$ and $X_3(c, c')$ that are constructed recursively in the tree T_X belong to the sets A_6, A_5 and A_2, respectively. Figure 7.7 illustrates how the extended regions $\mathcal{R}_{A_6}$, $\mathcal{R}_{A_5}$ and $\mathcal{R}_{A_2}$ match region R_X with respect to the three edges between v_X and its three neighbours in R_X, respectively. These three edges are the distinguished edges $e_{X_1(c,c')}$, $e_{X_2(c,c')}$ and $e_{X_3(c,c')}$. Thus, the element

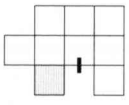


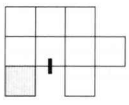
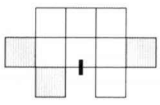
A_1 $(A_1, F_1, \{(A_2, A_1, A_2)\})$  $(A_1, F_2, \{(A_4, A_1, A_4)\})$  $(A_1, F_4, \{(A_6, A_5, A_2)\})$

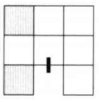
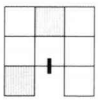
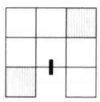
 $(A_1, F_7, \{(A_3, \emptyset, A_3)\})$  $(A_1, F_7, \{(A_6, \emptyset, A_3)\})$  $(A_1, F_6, \{(A_6, A_7, A_6)\})$

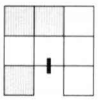
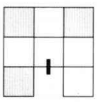
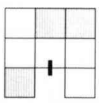


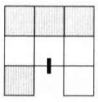
$(A_1, F_7, \{(A_6, \emptyset, A_6)\})$



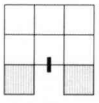
A_2 $(A_2, F_1, \{(A_5, A_1, A_2)\})$  $(A_2, F_1, \{(A_5, A_1, A_2)\})$  $(A_2, F_2, \{(A_7, A_1, A_4)\})$

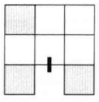
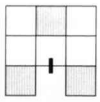
 $(A_2, F_4, \{(A_7, A_5, A_2)\})$  $(A_2, F_7, \{(A_6, \emptyset, A_3)\})$  $(A_2, F_4, \{(A_5, A_5, A_6)\})$

 $(A_2, F_7, \{(A_7, \emptyset, A_3)\})$  $(A_2, F_6, \{(A_7, A_7, A_6)\})$  $(A_2, F_7, \{(A_6, \emptyset, A_6)\})$



$(A_2, F_7, \{(A_7, \emptyset, A_7)\})$



A_3 $(A_3, F_2, \{(A_5, A_1, A_5)\})$  $(A_3, F_4, \{(A_7, A_5, A_5)\})$  $(A_3, F_7, \{(A_6, \emptyset, A_6)\})$

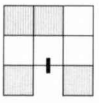
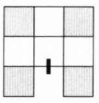
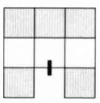
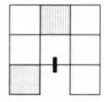
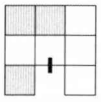
 $(A_3, F_7, \{(A_7, \emptyset, A_6)\})$  $(A_3, F_6, \{(A_7, A_7, A_7)\})$  $(A_3, F_7, \{(A_7, \emptyset, A_7)\})$

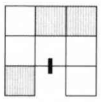
Figure 7.5. Elements in the $(\mathcal{A}, \mathcal{F})$-set (part 1 of 2).



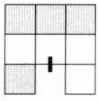
A_4 $(A_4, F_7, \{(A_6, \emptyset, A_3)\})$



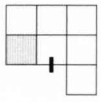
$(A_4, F_7, \{(A_7, \emptyset, A_3)\})$



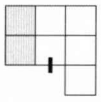
$(A_4, F_7, \{(A_6, \emptyset, A_6)\})$



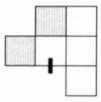
$(A_4, F_7, \{(A_7, \emptyset, A_6)\})$



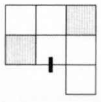
A_5 $(A_5, F_3, \{(\emptyset, A_2, A_2)\})$



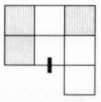
$(A_5, F_5, \{(\emptyset, A_5, A_2)\})$



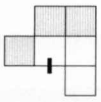
$(A_5, F_7, \{(\emptyset, \emptyset, A_3)\})$



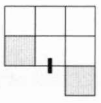
$(A_5, F_7, \{(\emptyset, A_6, A_6)\})$



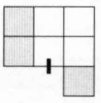
$(A_5, F_7, \{(\emptyset, A_7, A_6)\})$



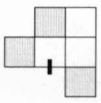
$(A_5, F_7, \{(\emptyset, \emptyset, A_6)\})$



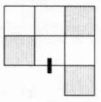
A_6 $(A_6, F_3, \{(\emptyset, A_2, A_5)\})$



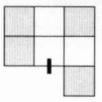
$(A_6, F_5, \{(\emptyset, A_5, A_5)\})$



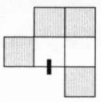
$(A_6, F_7, \{(\emptyset, \emptyset, A_6)\})$



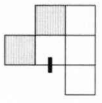
$(A_6, F_7, \{(\emptyset, A_6, A_7)\})$



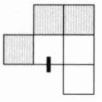
$(A_6, F_7, \{(\emptyset, A_7A_7)\})$



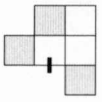
$(A_6, F_7, \{(\emptyset, \emptyset, A_7)\})$



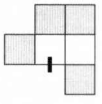
A_7 $(A_7, F_7, \{(\emptyset, \emptyset, A_3)\})$



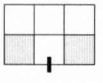
$(A_7, F_7, \{(\emptyset, \emptyset, A_6)\})$



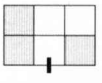
$(A_7, F_7, \{(\emptyset, \emptyset, A_6)\})$



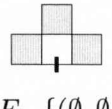
$(A_7, F_7, \{(\emptyset, \emptyset, A_7)\})$



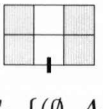
$(A_7, F_7, \{(\emptyset, A_3, \emptyset)\})$



$(A_7, F_7, \{(\emptyset, A_6, \emptyset)\})$



$(A_7, F_8, \{(\emptyset, \emptyset, \emptyset)\})$



$(A_7, F_7, \{(\emptyset, A_7, \emptyset)\})$

Figure 7.6. Elements in the $(\mathcal{A}, \mathcal{F})$-set (part 2 of 2).

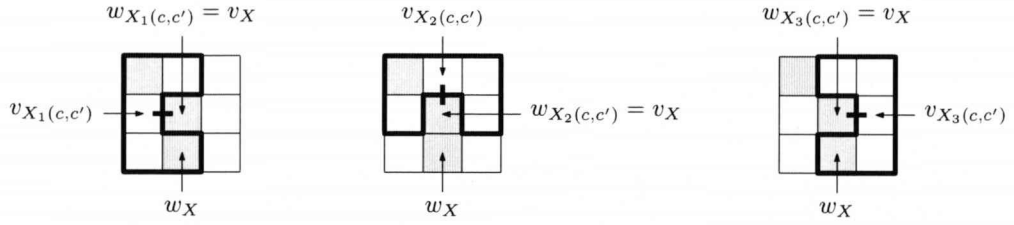


Figure 7.7. An example of a region of an edge-boundary pair $X \in A_1$. The extended regions $\mathcal{R}_{A_6}$, $\mathcal{R}_{A_5}$ and $\mathcal{R}_{A_2}$ match region R_X with respect to the three edges between v_X and its neighbours in R_X. Thus, the recursively constructed edge-boundary pair $X_1(c, c') \in A_6$, $X_2(c, c') \in A_5$ and $X_3(c, c') \in A_2$.

$(A_1, F_4, \{(A_6, A_5, A_2)\})$ is in S, as stated in Figure 7.5 below the extended region $\mathcal{R}$.

7.3 Strong spatial mixing and rapid mixing

In order to show that $\Gamma^d(X)$ is decreasing exponentially for every edge-boundary pair X, we verify that the $(\mathcal{A}, \mathcal{F})$-set S is good with respect to the values $\mu_{F_1}, \ldots, \mu_{F_8}$ in Lemma 7.1–7.4. For $q = 6$ we have the following lemma.

Lemma 7.5. *Suppose $q = 6$ and $\lambda \in [0, 1]$. The $(\mathcal{A}, \mathcal{F})$-set S is good with respect to $\varepsilon = 1/1000$ and the values $\mu_{F_1}, \ldots, \mu_{F_8}$ defined in Lemma 7.1.*

Proof. From Lemma 7.1 we have that μ_{F_i} is an upper bound on $\mu(X)$ for all edge-boundary pairs $X \in F_i$, $i = 1, \ldots, 8$. Let $\alpha_{A_1} = 100$, $\alpha_{A_2} = \alpha_{A_3} = 92$ and $\alpha_{A_4} = \alpha_{A_5} = \alpha_{A_6} = \alpha_{A_7} = 70$. For every element $(A, F, \{(A_{h_1}, A_{h_2}, A_{h_3})\}) \in S$ we successfully verify that $\mu_F(\alpha_{A_{h_1}} + \alpha_{A_{h_2}} + \alpha_{A_{h_3}}) \leq \alpha_A(1 - \varepsilon)$.

We make the following remark. We use the mathematics software Mathematica in order to verify that all inequalities are satisfied. The Mathematica file, including comments, can be accessed on the webpage `http://www.csc.liv.ac.uk/~markus/phdthesis/` □

Similarly, for $q = 5$, 4 and 3, we show that S is good with respect to $\mu_{F_1}, \ldots, \mu_{F_8}$ in Lemma 7.2, 7.3 and 7.4, respectively. Each of these three cases requires their own set of constants $\alpha_{A_1}, \ldots, \alpha_{A_7}$ that satisfy the inequalities in Definition 6.11. We have the following lemmas.

Lemma 7.6. *Suppose $q = 5$ and $\lambda \in [0.127, 0.286]$. The $(\mathcal{A}, \mathcal{F})$-set S is good with respect to $\varepsilon = 1/1000$ and the values $\mu_{F_1}, \ldots, \mu_{F_8}$ defined in Lemma 7.2.*

Lemma 7.7. *Suppose $q = 4$ and $\lambda \in [0.262, 0.373]$. The $(\mathcal{A}, \mathcal{F})$-set S is good with respect to $\varepsilon = 1/1000$ and the values $\mu_{F_1}, \ldots, \mu_{F_8}$ defined in Lemma 7.3.*

Lemma 7.8. *Suppose $q = 3$ and $\lambda \in [0.393, 0.474]$. The $(\mathcal{A}, \mathcal{F})$-set S is good with respect to $\varepsilon = 1/1000$ and the values $\mu_{F_1}, \ldots, \mu_{F_8}$ defined in Lemma 7.4.*

We are now ready to prove Theorem 3.2, which is restated below.

Theorem (3.2). *The anti-ferromagnetic Potts model on $\mathbb{Z}^2$ with parameters q and $\lambda \leq 1$, has strong spatial mixing in the following cases.*

(i) $q \geq 6$ *and* $\lambda \geq 0$,

(ii) $q = 5$ *and* $\lambda \geq 0.127$,

(iii) $q = 4$ *and* $\lambda \geq 0.262$, *and*

(iv) $q = 3$ *and* $\lambda \geq 0.393$.

Furthermore, for these values of q and λ, and any region R of $\mathbb{Z}^2$, the Glauber dynamics is rapidly mixing for any q_0-colouring $\mathcal{B}$ of ∂R. The mixing time $\tau(\delta) \in O(n^2 + n \log \frac{1}{\delta})$, where n is the number of vertices in R.

Proof. Let $\varepsilon = 1/1000$. From Lemma 7.5–7.8 and Lemma 6.12 it follows that there is a constant $\alpha \geq 0$ such that $\Gamma^d(X) \leq \alpha(1-\varepsilon)^d$ for every edge-boundary pair X, for $q = 6$, $\lambda \in [0, 1]$, and $q = 5$, $\lambda \in [0.127, 0.286]$, and $q = 4$, $\lambda \in [0.262, 0.373]$, and $q = 3$, $\lambda \in [0.393, 0.474]$. From Lemma 2.21 we have that the square lattice is uniformly sub-exponential. Using Lemma 6.3 gives strong spatial mixing and the mixing time of the Glauber dynamics. Combining the mixing bounds with Theorem 3.9 finishes the proof. □

7.4 The computational part

Lemma 7.1–7.4 are based on computation. Carrying out the computations by hand is not feasible since we have to consider millions of edge-boundary pairs X, and for each edge-boundary pair we will have to compute $\mu(X)$. We have implemented programs in C in order to compute the maximum value of $\mu(X)$ over all $X \in F_i$, $i = 1, \ldots, 8$. We chose C for this task because it is fast. The programs use only standard libraries. Even though the sizes of the regions F_i are rather small, and the number of colours q is not very high, we still do not want to implement the task in a brute-force manner. This would result in programs which could take days, maybe weeks, or even months, to run on a powerful PC.

The next sections describe how we deal with certain issues in order to keep the running time sufficiently short. Some of the ideas described here are also used in the computational part of the proofs of exponential decay for the triangular lattice (Chapter 8) and the kagome lattice (Chapter 9). Details on the implementation are referred to the webpage `http://www.csc.liv.ac.uk/~markus/phdthesis/`

We make the remark that all calculations carried out by the programs are exact in the sense that no "digits are lost" due to floating-point arithmetic. That is, all

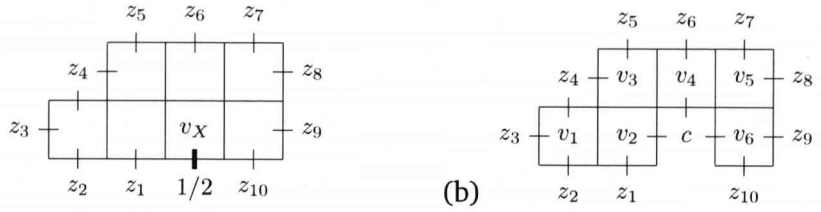


Figure 7.8. (a) A labelling of the boundary edges of an edge-boundary pair. (b) A labelling of the vertices used in the dynamic programming when computing the weights ω_c.

variables are integers, no divisions take place and a rational number is stored with its numerator and denominator.

7.4.1 Looping through colourings of the edge boundary

Let F be the set of all edge-boundary pairs X such that the region R_X is identical to some extended region $\mathcal{R}$ and e_X is the designated edge in $\mathcal{R}$. We wish to find the maximum value of $\mu(X)$ over all edge-boundary pairs X in F. There are many different edge-boundary pairs in F that share the same value $\mu(X)$, so in order to maximise $\mu(X)$ we would like to consider as few edge-boundary pairs as possible in F, but still make sure we cover one with the largest $\mu(X)$.

The edge-boundary pairs in F differ only on the colourings of the edge boundary. Therefore, edge-boundary pairs in F are considered by looping through colourings of the edge boundary. In order to cut down the search space, we consider, without loss of generality with respect to $\mu(X)$, only edge-boundary pairs X with the following constraints. We assume $B_X(e_X) = 1$ and $B'_X(e_X) = 2$. Let $z_1, \ldots, z_k$ be the colours of the boundary edges $\mathcal{E}R_X \backslash \{e_X\}$ in B_X, taken in any order. The colour of the first edge, z_1, is restricted to 1, 2 and 3 only. For every $i \in \{2, \ldots, k\}$, at least one colour in $\{z_1, \ldots, z_{i-1}\}$ is $z_i - 1$ if $z_i \geq 4$. That is, when considering the colours $z_1, \ldots, z_k$ in order, we introduce a next colour $c \geq 4$ only if all colours in $\{3, \ldots, c-1\}$ have been used on preceding edges. Note that any pair of colourings (B_X, B'_X) can be transformed into a pair of colourings satisfying these constraints, only by permuting the q colours. This will not change the value of $\mu(X)$.

We cut down the search space further by considering a defined ordering of the colours on boundary edges incident to "corner vertices". For $j > i$, let z_i and z_j be the colours on two boundary edges that are incident to the same vertex $v \in R_X$. If z_i and z_j defines the colour on only one boundary edge each, then we require $z_j \geq z_i$. For instance, see Figure 7.8(a). Here the ordering of the colours z_2 and z_3 is irrelevant in terms of the value $\mu(X)$. Therefore we require $z_3 \geq z_2$. The same holds for z_7 and z_8, and z_9 and z_{10}. Note that colour z_4 is defining the colour on two boundary edges.

7.4.2 Computing $\mu(X)$

For each edge-boundary pair X we consider, we have to compute $\mu(X)$. In fact, due to the symmetry between colour 1 and 2 in the previous section, we want to compute $\mu(X)$ as $\mu_{1,2}(X)$. Since we need to compute an upper bound on $\mu(X)$ over an interval of λ, we want to derive $\mu(X)$ as a rational function in λ. We do this using a dynamic programming approach, which is demonstrated with the following example.

Consider the edge-boundary pair in Figure 7.8(a). For $c \in Q$, we derive the weight ω_c as follows. Assign colour c to vertex v_X and let $v_1, \ldots, v_6$ be the other vertices in R_X according to Figure 7.8(b). For vertex v_i, let $\omega_j(i)$ be the weight contributed by the subregion consisting of vertices $v_1, \ldots, v_i$, such that v_i has colour j. Let $W(i) = \sum_{j=1}^{q} \omega_j(i)$, and hence $\omega_c = W(6)$. Let $n_j(i)$ be the number of boundary edges with colour j that are incident to v_i. So, for vertex v_6, the colours z_9, z_{10} and c count towards the values $n_j(6)$. Starting with vertex v_1, we have $\omega_j(1) = \lambda^{n_j(1)}$ for $j = 1, \ldots, q$. For $i \geq 2$, let

$$\omega'_j(i) = W(i-1) - \omega_j(i-1) + \lambda\omega_j(i-1),$$

where λ in the last term is the contribution from the monochromatic edge between v_{i-1} and v_i when they both have colour j. We now adjust for monochromatic boundary edges, which gives us $\omega_j(i) = \omega'_j(i) \cdot \lambda^{n_j(i)}$.

When implementing the dynamic programming procedure we keep track of the coefficients in the polynomials $\omega_j(i)$.

Chapter 8

The triangular lattice

In this chapter we prove Theorem 3.3. We define a collection $\mathcal{A}$ of 2048 sets of edge-boundary pairs, and a collection $\mathcal{F}$ of 39 sets of edge-boundary pairs. We construct a valid $(\mathcal{A}, \mathcal{F})$-set S which contains 126,967 elements and is therefore generated with computer assistance. For every set $F \in \mathcal{F}$ we compute a value μ_F which is the maximum value of $\mu(X)$ over all edge-boundary pairs $X \in F$. These values are obtained with the help of a computer, and it is an important task to keep the running time short. Here short means that we are able to obtain the values in a matter of two weeks. We show that the set S is good with respect to the 39 values μ_F, which implies exponential decay of $\Gamma^d(X)$ for every edge-boundary pair X.

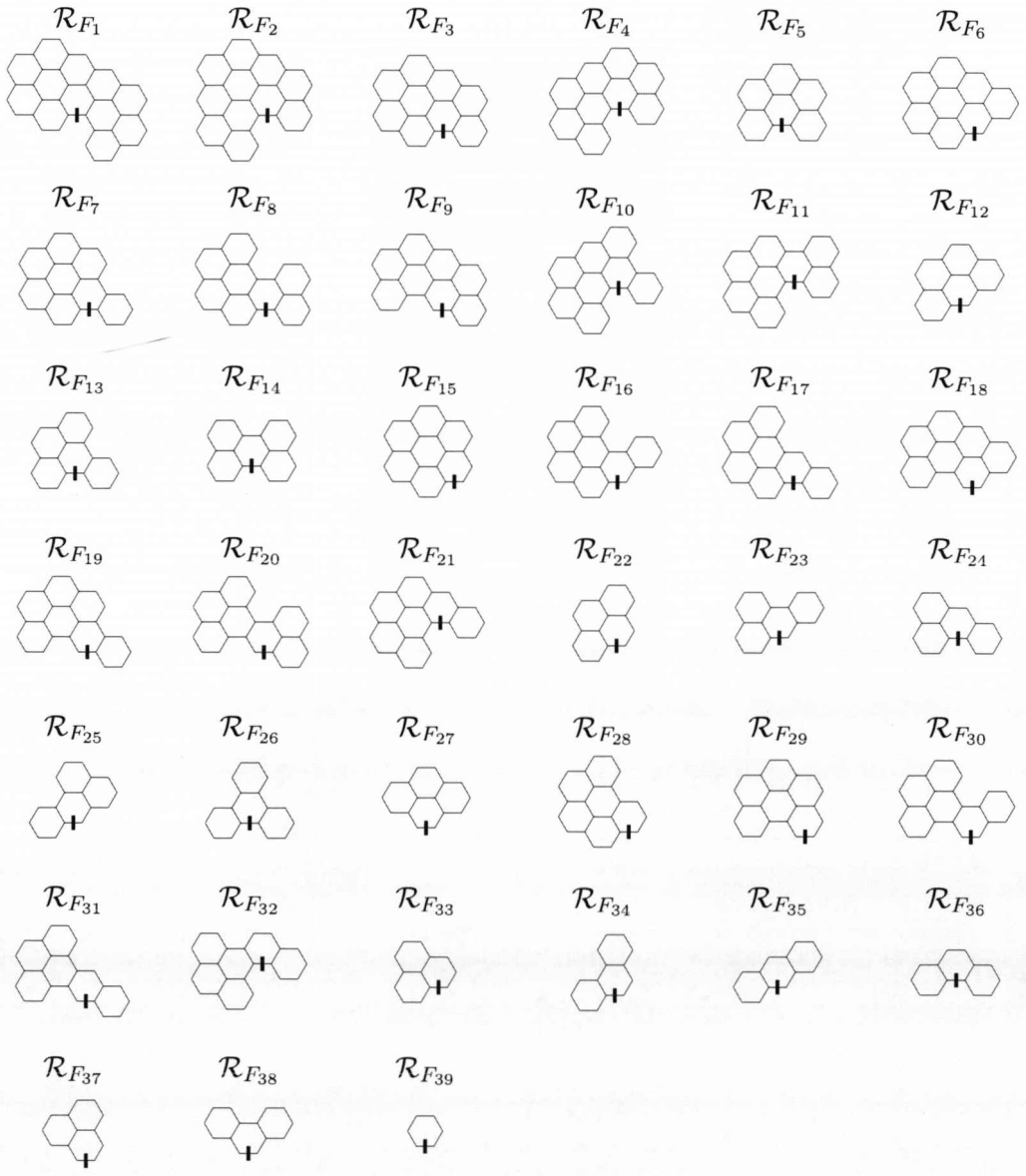
The organisation of the chapter is as follows. In Section 8.1 we define the collections $\mathcal{A}$ and $\mathcal{F}$ of edge-boundary pairs and present the 39 values μ_F. We explain how the $(\mathcal{A}, \mathcal{F})$-set S is constructed and give the proof of Theorem 3.3. In Section 8.2 we explain how we tackled the main difficulties with the implementation of the computational part of the proof.

8.1 Exponential decay

For $i = 1, \ldots, 39$, let $\mathcal{R}_{F_i}$ be the extended region illustrated in Figure 8.1, and let F_i be the set of edge-boundary pairs X such that R_X matches $\mathcal{R}_{F_i}$ with respect to e_X. Let $\mathcal{F} = \{F_1, \ldots, F_{39}\}$. The following lemma is proved with the help of a computer.

Lemma 8.1. *Let $q = 9$ and $\lambda = 0$. For $i = 1, \ldots, 39$ and any edge-boundary pair $X \in F_i$, we have $\mu(X) \leq \mu_{F_i}$, where the value of μ_{F_i} is given in Figure 8.1.*

Like Lemmas 7.1–7.4 for the square lattice, we prove the lemma by applying Corollary 6.6. That is, in order to obtain the value μ_{F_i}, we consider every edge-boundary pair X with region $R_X = \mathcal{R}_{F_i}$. The value μ_{F_i} is obtained by maximising $\mu(X)$ over all colourings of the boundary of the region $\mathcal{R}_{F_i}$. Details on the computational part of the proof will be discussed in Section 8.2.



$\mu_{F_1} = 68809973/310505657$
$\mu_{F_2} = 11623551/51797443$
$\mu_{F_3} = 456459/2005687$
$\mu_{F_4} = 408609/1601573$
$\mu_{F_5} = 33/127$
$\mu_{F_6} = 18199/78779$
$\mu_{F_7} = 75312/325193$
$\mu_{F_8} = 14165/58613$
$\mu_{F_9} = 70661/293514$
$\mu_{F_{10}} = 31648/123341$
$\mu_{F_{11}} = 2655/10063$
$\mu_{F_{12}} = 521/1853$
$\mu_{F_{13}} = 208/757$
$\mu_{F_{14}} = 25/91$
$\mu_{F_{15}} = 9334/40215$
$\mu_{F_{16}} = 11332/46633$
$\mu_{F_{17}} = 11332/46633$
$\mu_{F_{18}} = 7067/29188$
$\mu_{F_{19}} = 11332/46633$
$\mu_{F_{20}} = 775/2941$
$\mu_{F_{21}} = 4248/16015$
$\mu_{F_{22}} = 21/73$
$\mu_{F_{23}} = 5/17$
$\mu_{F_{24}} = 5/17$
$\mu_{F_{25}} = 5/17$
$\mu_{F_{26}} = 32/113$
$\mu_{F_{27}} = 21/73$
$\mu_{F_{28}} = 2833/11551$
$\mu_{F_{29}} = 2833/11551$
$\mu_{F_{30}} = 620/2321$
$\mu_{F_{31}} = 620/2321$
$\mu_{F_{32}} = 688/2389$
$\mu_{F_{33}} = 5/17$
$\mu_{F_{34}} = 4/13$
$\mu_{F_{35}} = 4/13$
$\mu_{F_{36}} = 4/13$
$\mu_{F_{37}} = 5/17$
$\mu_{F_{38}} = 4/13$
$\mu_{F_{39}} = 1/3$

Figure 8.1. Extended regions we use when defining the collection $\mathcal{F}$ of sets of edge-boundary pairs. Below are the values μ_{F_i}.

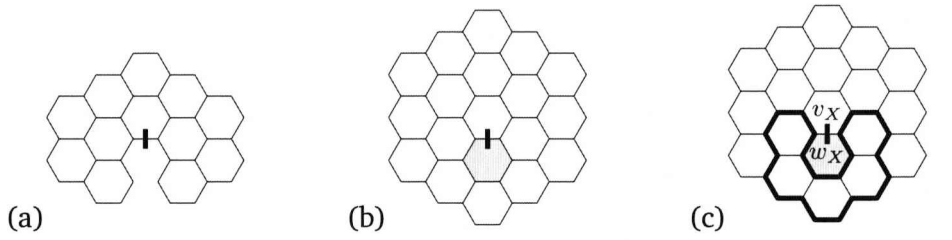


Figure 8.2. (a) The extended region $\mathcal{R}_A$. (b) The extended region $\mathcal{R}_{\text{big}}$ with all vertices labelled "in". Note that the shaded vertex is not a part of the region. (c) The vertex v_X is always labelled "in" and at least one of the five vertices within the thick border is labelled "out". Hence there are $2^{12}(2^5 - 1) = 126,967$ combinations of "in"- and "out"-labels of the vertices of $\mathcal{R}_{\text{big}}$ that match an edge-boundary pair X.

Let $\mathcal{R}_A$ be the extended region in Figure 8.2(a), consisting of 12 vertices. There are $2^{11} = 2048$ different combinations of labelling the vertices in $\mathcal{R}_A$ "in" and "out", such that the vertex that is incident to the designated edge is labelled "in". We define $\mathcal{R}_{A_1}, \ldots, \mathcal{R}_{A_{2048}}$ to be all those different extended regions. For $i = 1, \ldots, 2048$, let A_i be the set of edge-boundary pairs X such that R_X matches $\mathcal{R}_{A_i}$ with respect to edge e_X, and let $\mathcal{A} = \{A_1, \ldots, A_{2048}\}$.

We now define a valid $(\mathcal{A}, \mathcal{F})$-set S. The set S contains 126,967 elements and is therefore constructed with the help of a computer. Let $\mathcal{R}_{\text{big}}$ be the extended region in Figure 8.2(b). For any edge-boundary pair X, the region R_X will match $\mathcal{R}_{\text{big}}$ with respect to e_X for a suitable combination of the labels "in" and "out" on the vertices of $\mathcal{R}_{\text{big}}$. The vertex v_X is always in R_X and at least one of the vertices adjacent to w_X is not in R_X (recall Definition 5.2 of an edge-boundary pair). That is, at least one of the five vertices within the thick border in Figure 8.2(c) has to be excluded from R_X. There are $2^5 - 1$ ways of assigning labels "in" and "out" to these five vertices such that at least one is "out". With vertex v_X always being "in", there are 12 remaining vertices that are either "in" or "out". Hence there are $2^{12}(2^5 - 1) = 126,967$ combinations of "in"- and "out"-labels of the vertices of $\mathcal{R}_{\text{big}}$ that match an edge-boundary pair.

Fix an edge-boundary pair X and suppose that $\mathcal{R}_{\text{big}}$ (with some "in" and "out" labels) is the extended region matching X. Now, let $A \in \mathcal{A}$ be the set of edge-boundary pairs such that $\mathcal{R}_A$ matches R_X with respect to e_X, and hence $X \in A$. It follows that $\mathcal{R}_A$ is the subregion of $\mathcal{R}_{\text{big}}$ that is outlined with a thick border in Figure 8.3(a).

For any pair of distinct colours $c, c' \in Q$, and $i \in \{1, \ldots, 5\}$, let $X_i(c, c')$ be the edge-boundary pair that is constructed recursively in the tree T_X, where i is the index in the clockwise ordering of the neighbours of v_X, starting with the vertex after w_X. If the ith neighbour of v_X is not in R_X, then let $h_i = 0$, otherwise let $h_i \in \{1, \ldots, 2048\}$ be the index such that $\mathcal{R}_{A_{h_i}}$ is one of the subregions of $\mathcal{R}_{\text{big}}$ in Figure 8.3(b)–(f), depending on which neighbour of v_X is considered. Note that the designated edge of $\mathcal{R}_{A_{h_i}}$ coincides with the edge between v_X and its ith neighbour. It follows that the edge-boundary

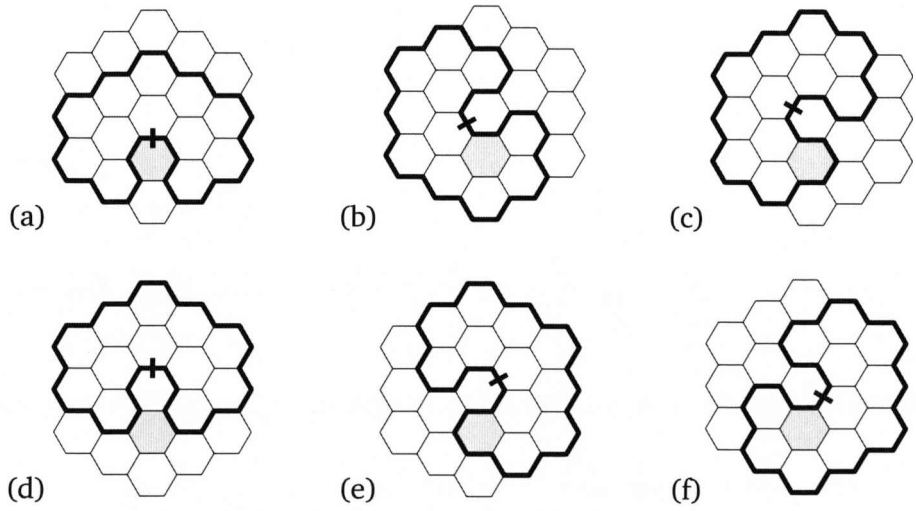


Figure 8.3. The extended region $\mathcal{R}_A$, outlined with a thick border, as a subregion of $\mathcal{R}_{\text{big}}$, taken such that the designated edge of $\mathcal{R}_A$ coincides with different edges of $\mathcal{R}_{\text{big}}$.

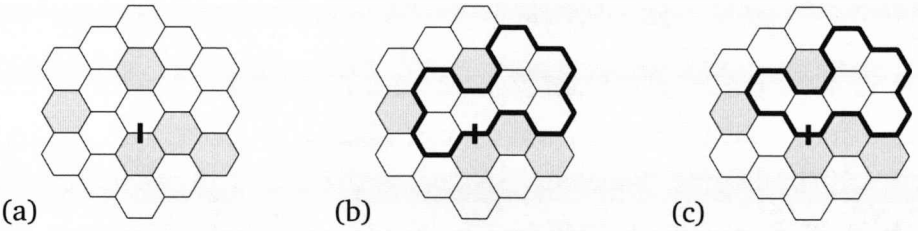


Figure 8.4. An example of an extended region $\mathcal{R}_{\text{big}}$ (a). The extended regions $\mathcal{R}_{F_{20}}$ (b) and $\mathcal{R}_{F_{30}}$ (c), outlined with a thick border, are subregions of $\mathcal{R}_{\text{big}}$ such that all vertices are labelled "in" in $\mathcal{R}_{\text{big}}$.

pair $X_i(c, c')$ belongs to the set A_{h_i}, provided the ith neighbour of v_X is in R_X. That is, $R_{X_i(c,c')}$ is matching $\mathcal{R}_{A_{h_i}}$ with respect to $e_{X_i(c,c')}$. Thus, the 5-tuple $(A_{h_1}, \ldots, A_{h_5})$ represents the sets $A_{h_1}, \ldots, A_{h_5}$ to which the recursively constructed edge-boundary pairs $X_1(c, c'), \ldots, X_5(c, c')$ belong. We define $A_0 = \emptyset$.

Let $\mathcal{F}' \subseteq \mathcal{F}$ contain the sets F in $\mathcal{F}$, such that $X \in F$. That is, $F \in \mathcal{F}$ is in $\mathcal{F}'$ if $\mathcal{R}_F$ is a subregion of $\mathcal{R}_{\text{big}}$ such that the designated edge in $\mathcal{R}_F$ and $\mathcal{R}_{\text{big}}$ coincide, and all vertices of $\mathcal{R}_F$ are labelled "in" in $\mathcal{R}_{\text{big}}$. Let F_{min} be the set in $\mathcal{F}'$ that minimises μ_F over all $F \in \mathcal{F}'$. From Lemma 8.1 we have that $\mu(X) \leq \mu_{F_{\text{min}}}$. For example, consider $\mathcal{R}_{\text{big}}$ in Figure 8.4(a). In Figure 8.4(b) we see that $\mathcal{R}_{F_{20}}$ is a subregion of $\mathcal{R}_{\text{big}}$ for which all vertices are labelled "in". Similarly, in Figure 8.4(c) we see that $\mathcal{R}_{F_{30}}$ is also a subregion of $\mathcal{R}_{\text{big}}$ for which all vertices are labelled "in". Hence F_{20} and F_{30} are in $\mathcal{F}'$. There are more sets in $\mathcal{F}'$, but the the set $F \in \mathcal{F}'$ that minimises μ_F is F_{20}. Thus, $F_{\text{min}} = F_{20}$.

The procedure above describes how an element $(A, F_{\text{min}}, \{(A_{h_1}, \ldots, A_{h_5})\})$ in

the $(\mathcal{A}, \mathcal{F})$-set S is uniquely derived from an extended region $\mathcal{R}_{\text{big}}$. The procedure is implemented as a computer program, which is generating the whole set S by looping through all 126,967 extended regions $\mathcal{R}_{\text{big}}$. From the construction of S it follows that S is valid. The next lemma, which is proved by computer assistance, says that S is good with respect to the values $\mu_{F_1}, \ldots, \mu_{F_{39}}$ in Figure 8.1. See Section 8.2 for details on the computational part of the proof.

Lemma 8.2. *The $(\mathcal{A}, \mathcal{F})$-set S is good with respect to $\varepsilon = 1/1000$ and the 39 values $\mu_{F_1}, \ldots, \mu_{F_{39}}$ in Figure 8.1.*

Proof. Lemma 8.1 says that $\max_{X \in F_i} \mu(X) \leq \mu_{F_i}$ for $i = 1, \ldots, 39$. For every $A \in \mathcal{A}$ there is a constant $1.1 \leq \alpha_A \leq 5$. The exact value of α_A can be found on the webpage
`http://www.csc.liv.ac.uk/~markus/phdthesis/`
By using a computer, we successfully verify for each element $(A, F, \{(A_{h_1}, \ldots, A_{h_5})\})$ in S that $\mu_F(\alpha_{A_{h_1}} + \cdots + \alpha_{A_{h_5}}) \leq \alpha_A(1 - \varepsilon)$. □

Next is the proof of Theorem 3.3. First we restate the theorem.

Theorem (3.3). *The anti-ferromagnetic Potts model on the triangular lattice with parameters $q = 9$ and $\lambda = 0$ has strong spatial mixing, and the Glauber dynamics is rapidly mixing for any region R and any q_0-colouring $\mathcal{B}$ of ∂R. The mixing time $\tau(\delta) \in O(n^2 + n \log \frac{1}{\delta})$, where n is the number of vertices in R.*

Proof. From Lemma 8.2 and Lemma 6.12 it follows that there is a constant $\alpha \geq 0$ such that $\Gamma^d(X) \leq \alpha(1 - \varepsilon)^d$ for every edge-boundary pair X, where $\varepsilon = 1/1000$. From Lemma 2.22 we have that the triangular lattice is uniformly sub-exponential. Strong spatial mixing and the mixing time of the Glauber dynamics now follow from Lemma 6.3. □

8.2 The computational part

The computational part of the proof of exponential decay consists of two tasks: computing the values $\mu_{F_1}, \ldots, \mu_{F_{39}}$ in Figure 8.1, and constructing the $(\mathcal{A}, \mathcal{F})$-set S and verifying that it is good with respect to $\mu_{F_1}, \ldots, \mu_{F_{39}}$.

8.2.1 Verifying that the $(\mathcal{A}, \mathcal{F})$-set S is good

Due to the large number of inequalities, we use a computer to find values of the variables $\alpha_{A_1}, \ldots, \alpha_{A_{2048}}$ that satisfy them. As explained on page 80, at the end of Section 6.4, a solution to the set of inequalities can be obtained by solving a linear program. We have implemented a program in C, using only standard libraries, that follows the procedure described above. The program outputs a file which represents a linear

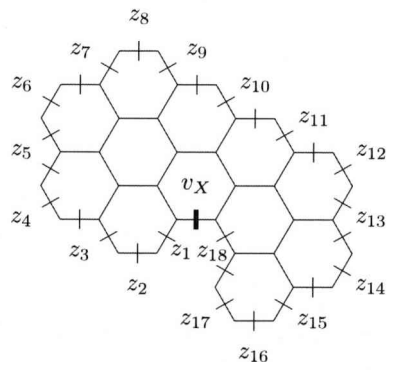


Figure 8.5. A labelling of the boundary edges of an edge-boundary pair.

program that contains all the inequalities that have to be satisfied in order to make S good with respect to $\mu_{F_1}, \ldots, \mu_{F_{39}}$. The linear program is specified in the CPLEX LP format which is a widely used format used by many linear-program solvers. We solve the linear program with the GNU Linear Programming Kit, which is a free large-scale linear-program solver. Doing this with the constant ε set to $1/1000$, we successfully satisfy all inequalities with $\alpha_{A_1}, \ldots, \alpha_{A_{2048}}$ being values in the interval $[1.1, 5]$. For details on these computations, see the webpage `http://www.csc.liv.ac.uk/~markus/phdthesis/`

8.2.2 Computing the values $\mu_{F_1}, \ldots, \mu_{F_{39}}$

In order to compute the values $\mu_{F_1}, \ldots, \mu_{F_{39}}$ in Figure 8.1, we loop through colourings of the edge boundaries of the regions in Figure 8.1 in the same manner as explained in Section 7.4. Unlike the regions we considered for the square lattice, some of the regions we deal with here have boundary edges incident to vertex w_X, where X is the edge-boundary pair we consider. These edges can only be coloured 1 and 2, supposing the edge e_X has colour 1 and 2 in B_X and B'_X, respectively. Since we are maximising $\mu(X)$ by computing the value $\mu_{1,2}(X)$, we have to consider combinations of the two colours on the edges incident to w_X. For example, the colours z_1 and z_{18} in Figure 8.5 are either 1 and 2, or 2 and 1, respectively. Note that, unlike the situations with the square lattice in the previous chapter, $\lambda = 0$ here. We do not have to deal with a range of λ as before. Instead of treating $\mu(X)$ as a polynomial in λ, $\mu(X)$ is here a rational number.

We have implemented programs in C that compute the 39 values $\mu_{F_1}, \ldots, \mu_{F_{39}}$. We chose C because it is fast. We only use standard libraries. Some of the 39 regions we consider are rather large. Also the number of colours $q = 9$ is a fairly large number. Therefore the running time of the program is somewhat long, despite implementation optimisations of the C-code. Examples of optimisations are that we try to avoid unnecessary memory allocation (we pass on pointers instead of duplicating large memory blocks when calling functions), and we avoid unnecessary traversal of

arrays (using pointers and linked lists when possible). On a fairly powerful home-PC as of year 2006, it takes a couple of weeks to obtain all values $\mu_{F_1}, \ldots, \mu_{F_{39}}$. Most of the running time is spent on only a few of the regions, in particular the ones with a large edge boundary. We left a computer running day and night. For details on the implementation, see webpage `http://www.csc.liv.ac.uk/~markus/phdthesis/`

8.2.3 The experimental phase

When taking on the task of showing exponential decay for the triangular lattice with nine colours, it was not clear how to define the collections $\mathcal{A}$ and $\mathcal{F}$ of sets of edge-boundary pairs. There has been a process of trial-and-error prior to the final definitions of $\mathcal{A}$ and $\mathcal{F}$. We started with a few regions that were not very large, and step by step we increased the size and the number of regions until we successfully could show exponential decay. During this experimental phase we had to compute μ_F for many more regions $\mathcal{R}_F$ than the 39 regions in Figure 8.1. Some of the regions we considered were even larger than the ones in the figure. Obtaining the exact values of μ_F would take weeks, or months of running time, per region. Furthermore, it would have been a tedious job writing the code and optimising it for every single region. Therefore, instead of exactly computing the maximum value of μ_F for a given region $\mathcal{R}_F$, we first computed the value approximately using a hill-climbing heuristic.

For a given region $\mathcal{R}_F$ we did not loop through all colourings of the boundary. Instead we considered edge-boundary pairs X with $R_X = \mathcal{R}_F$ and for which the colouring of the boundary was randomised. We maximised $\mu(X)$ over such randomised edge-boundary pairs X. The maximum value of $\mu(X)$ that we obtained could of course be less than the true maximum value, since we might have missed the "worst" colouring of the boundary. However, by maximising $\mu(X)$ over some randomised edge-boundary pairs enables us to obtain some value that could be used when constructing the $(\mathcal{A}, \mathcal{F})$-set and defining the inequalities in the linear program. In order to obtain a "maximum" value of $\mu(X)$, we repeatedly randomised boundaries until the largest value of $\mu(X)$ that was computed did not seem to get any larger. The total running time per region could be from a few minutes up to a few hours. The "maximum" value of $\mu(X)$ was often found within the first few minutes. However, for the biggest regions there could be a slight increase of the largest value after a couple of hours running time. When randomising colourings of the boundary it turned out to be efficient to make repeated random changes to the colouring of the boundary that had generated the largest value of $\mu(X)$ found so far. This approach let us obtain the "maximum" $\mu(X)$ much quicker than generating new colourings of the whole boundary at each step. During the experimental phase we built a library of heuristically approximated values of μ_F for a large set of regions $\mathcal{R}_F$.

The first solvable linear program that we obtained made use of many more regions $\mathcal{R}_F$ than the 39 regions in Figure 8.1. Since the exact value of μ_F had to be computed for each region $\mathcal{R}_F$, we first reduced the number of regions by replacing some of them with a suitable choice of subregion such that the linear program could still be solved. The outcome of this procedure resulted in the 39 regions in Figure 8.1, for which we exactly computed the maximum value μ_F as described in the previous section. We make the remark that the approximated values we had been using during the experimental phase were actually identical to the 39 true values $\mu_{F_1}, \ldots, \mu_{F_{39}}$.

Chapter 9

The kagome lattice

In this chapter we prove Theorem 3.4 and 3.5. We define a collection $\mathcal{A}$ which contains 684 sets of edge-boundary pairs, and a collection $\mathcal{F}$ which contains 9440 sets of edge-boundary pairs. Unlike in previous chapters, the collections $\mathcal{A}$ and $\mathcal{F}$ contain sets of edge-boundary pairs X that depend not only on the vertices in R_X but also on the colour of one of the boundary edges. Incorporating colours of boundary edges enables us to be more specific about the boundary colourings of the edge-boundary pairs that are constructed recursively in the tree T_X. For instance, Let X be an edge-boundary pair for which edge e_X has colour c in B_X and colour c' in B'_X, and for which $\mu_{c,c'}(X) \geq \mu_{c',c}(X)$. Let (C, C') be a pair of colourings of R_X drawn from $\Psi_X^{\min}$. If C and C' differ on v_X then it follows that v_X has colour c in C'. Hence, for any edge-boundary pair $X_i(\cdot,\cdot)$ that is constructed recursively in T_X, the colour of edge $e_{X_i(\cdot,\cdot)}$ is c in either $B_{X_i(\cdot,\cdot)}$ or $B'_{X_i(\cdot,\cdot)}$. Thus, if there is a boundary edge $e \in \mathcal{E}R_X$ that has colour c in X, then the colour of e in $X_i(\cdot,\cdot)$ must be the colour of $e_{X_i(\cdot,\cdot)}$ in either $B_{X_i(\cdot,\cdot)}$ or $B'_{X_i(\cdot,\cdot)}$. Restricting the possible colours of boundary edges of an edge-boundary pair X implies that the value $\mu(X)$ is likely to be lower than if maximised over all colourings of the boundary. As we have seen in the definition of a good $(\mathcal{A}, \mathcal{F})$-set, it is desirable to have small values of $\mu(X)$. In this chapter we introduce four sets $M_1, \ldots, M_4$ of edge-boundary pairs which capture the colour of one of the boundary edges. When defining the two collections $\mathcal{A}$ and $\mathcal{F}$ we combine these four sets with sets of edge-boundary pairs X that depend only on the vertices in R_X.

For each set $F \in \mathcal{F}$, we compute a value μ_F such that $\mu(X) \leq \mu_F$ for all edge-boundary pairs $X \in F$. The values μ_F are obtained with computer assistance, and we use certain tricks to keep the running time of the program short. When maximising $\mu(X)$ we note that some colourings of the boundary give very low values of $\mu(X)$. For instance, $\mu(X)$ will most likely not be maximised if all boundary edges are set to the same colour. When looping though colourings of the boundary we would like to avoid colourings that would surely not maximise $\mu(X)$. In order to filter out such colourings,

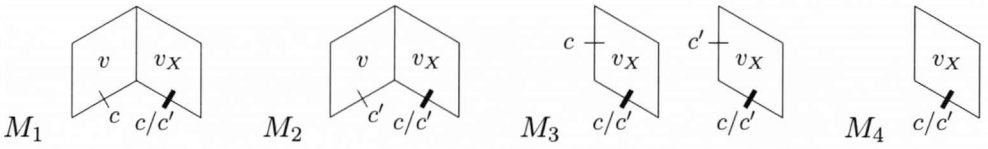


Figure 9.1. These diagrams illustrate how the sets $M_1, \ldots, M_4$ of edge-boundary pairs X are defined. Note that $\mu_{c,c'}(X) \geq \mu_{c',c}(X)$.

we split the region into two halves and consider each half separately. We throw away colourings for each half before combining them to obtain the value $\mu(X)$. As will be explained later, the colourings we throw away are guaranteed not to maximise $\mu(X)$. By splitting the region and filtering out colourings enable us to obtain all 9440 values μ_F in a matter of a couple of days running time. It would not be feasible to compute all values in a reasonable amount of time if we use the same approach as with the triangular lattice in the previous chapter. We construct a valid $(\mathcal{A}, \mathcal{F})$-set S and show that it is good with respect to the values μ_F, which implies exponential decay of $\Gamma^d(X)$ for every edge-boundary pair X. Due to the large size of the collections $\mathcal{A}$ and $\mathcal{F}$, we have to construct S with the help of a computer. In order to show that S is good, we solve a linear program containing around 160,000 inequalities and 684 unknowns.

We have to be careful when proving rapid mixing of the Glauber dynamics (Theorem 3.5) since the number of colours $q = 5$ is less than $\Delta + 2$. Hence we cannot use Lemma 6.3 to establish rapid mixing. We will deal with this in detail in Section 9.3.

The organisation of this chapter is as follows. In Section 9.1 we define the two collections $\mathcal{A}$ and $\mathcal{F}$. In Section 9.2 we define a valid $(\mathcal{A}, \mathcal{F})$-set and prove the strong spatial mixing result in Theorem 3.4. In Section 9.3 we prove that the Glauber dynamics is rapidly mixing. Here we present the proof of Theorem 3.5. Lastly, in Section 9.4 we discuss the method of splitting regions when computing the values μ_F.

9.1 Defining $\mathcal{A}$ and $\mathcal{F}$

We define four sets $M_1, \ldots, M_4$ of edge-boundary pairs. Let X be any edge-boundary pair and suppose that the colour of edge e_X is c in B_X and c' in B'_X, and suppose the weight $\omega_c \geq \omega_{c'}$. Hence $\mu(X) = \mu_{c,c'}(X) \geq \mu_{c',c}(X)$. Let v be the vertex that is a neighbour to both v_X and w_X. The sets $M_1, \ldots, M_4$ are defined as follows (see Figure 9.1).

- M_1 contains X if $v \in R_X$ and the colour of edge (w_X, v) is c.
- M_2 contains X if $v \in R_X$ and the colour of edge (w_X, v) is c'.
- M_3 contains X if $v \notin R_X$ and the colour of edge (v, v_X) is c or c'.
- M_4 contains X if $v \notin R_X$.

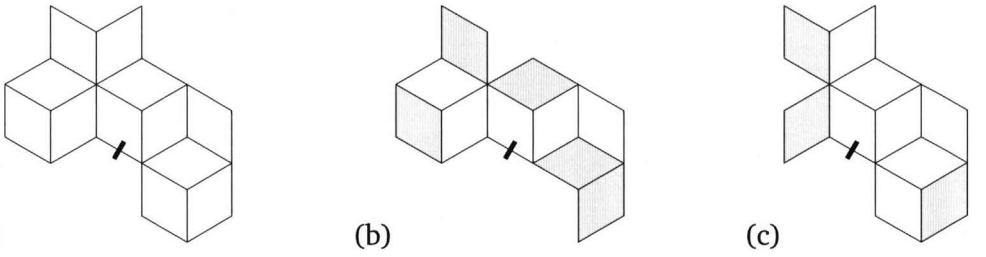


Figure 9.2. (a) The extended region $\mathcal{R}_A$. (a)–(c) Examples of three of the extended regions $\mathcal{R}_{A_1}, \ldots, \mathcal{R}_{A_{342}}$. The region R_X of any edge-boundary pair X matches exactly one of the extended regions $\mathcal{R}_{A_1}, \ldots, \mathcal{R}_{A_{342}}$ with respect to e_X.

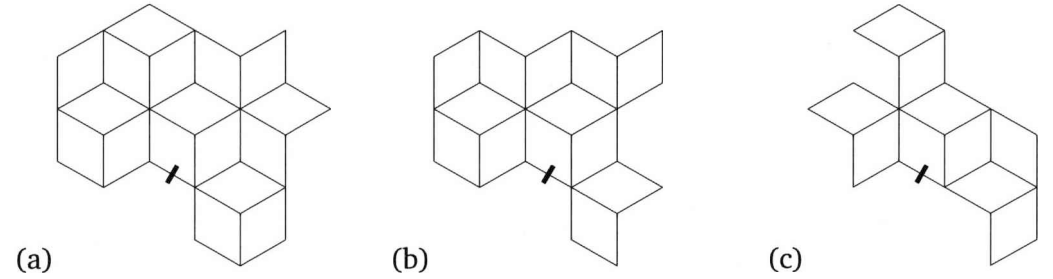


Figure 9.3. (a) The extended region $\mathcal{R}_F$. (a)–(c) Examples of three of the extended regions $\mathcal{R}_{F_1}, \ldots, \mathcal{R}_{F_{4720}}$. Note that all vertices in these extended regions are labelled "in".

Let $\mathcal{R}_A$ be the extended region in Figure 9.2(a). For $i = 1, \ldots, 342$, we define the extended regions $\mathcal{R}_{A_i}$ to be subregions of $\mathcal{R}_A$. These extended regions are defined such that for any edge-boundary pair X, the region R_X matches exactly one of them with respect to edge e_X. We are not explicitly stating all 342 extended regions here, however, as an example, the extended regions in Figure 9.2(a), (b) and (c) illustrate three of the extended regions $\mathcal{R}_{A_1}, \ldots, \mathcal{R}_{A_{342}}$. For $a = 1, \ldots, 342$, let A_a be the set of edge-boundary pairs X such that R_X matches $\mathcal{R}_{A_a}$ with respect to edge e_X. Furthermore, for $m = 1, \ldots, 4$ we define $A_{a,m} = A_a \cap M_m$, and define $\mathcal{A}$ to be the collection of all sets $A_{a,m}$ such that $A_{a,m} \neq \emptyset$. From the definition of $M_1, \ldots, M_4$ it follows that $\mathcal{A}$ contains 684 sets.

Let $\mathcal{R}_F$ be the extended region in Figure 9.3(a). The 4720 extended regions $\mathcal{R}_{F_1}, \ldots, \mathcal{R}_{F_{4720}}$ are all defined to be subregions of $\mathcal{R}_F$. All vertices are labelled "in" for each of these extended regions. Figures 9.3(a), (b) and (c) illustrate three examples of the extended regions $\mathcal{R}_{F_1}, \ldots, \mathcal{R}_{F_{4720}}$. For $f = 1, \ldots, 4720$, let F_f be the set of edge-boundary pairs X such that R_X matches $\mathcal{R}_{F_f}$ with respect to edge e_X. Furthermore, for $m = 1, \ldots, 4$ we define $F_{f,m} = F_f \cap M_m$, and define $\mathcal{F}$ to be the collection of all sets $F_{f,m}$ such that $F_{f,m} \neq \emptyset$. It follows that $\mathcal{F}$ contains 9440 sets. One of the extended regions $\mathcal{R}_{F_1}, \ldots, \mathcal{R}_{F_{4720}}$ is defined to contain only the single vertex that is incident to the designated edge. Therefore, any edge-boundary pair X is guaranteed to belong to at least one of the sets in $\mathcal{F}$. The following lemma is proved with computer assistance.

Lemma 9.1. *Let $q = 5$ and $\lambda = 0$. For every $f \in \{1, \ldots, 4720\}$ and $m \in \{1, \ldots, 4\}$ there is a constant $\mu_{F_{f,m}}$ such that for any edge-boundary pair $X \in F_{f,m}$, we have $\mu(X) \leq \mu_{F_{f,m}}$. The constants can be found on the webpage* `http://www.csc.liv.ac.uk/~markus/phdthesis/`

Unlike the similar lemmas for the square lattice and the triangular lattice in the previous chapters, we do not prove this lemma by applying Corollary 6.6. The reason for this is that we include information about colours of boundary edges in the sets $F_{f,m}$. Instead of using Corollary 6.6 we use Lemma 6.5. In order to obtain the value $\mu_{F_{f,m}}$, we consider every edge-boundary pair X with region $R_X = \mathcal{R}_{F_f}$. The value $\mu_{F_{f,m}}$ is obtained by maximising $\mu_{c,c'}(X)$ over all colourings (B_X, B'_X) of the boundary of the region $\mathcal{R}_{F_f}$ such that $B_X(e_X) = c$ and $B'_X(e_X) = c'$ and X is in set M_m. Corollary 6.6 would upper-bound $\mu(X)$, but here we are more precise and upper-bound $\mu_{c,c'}(X)$ under the constraint that X is in M_m. If we would use Corollary 6.6 to maximise $\mu(X)$ then we would have to consider both $\mu_{c,c'}(X)$ and $\mu_{c',c}(X)$, or equivalently, disregard the constraint that X belongs to M_m. This would result in larger values $\mu_{F_{f,m}}$ since removing the constraint gives us more freedom to choose boundary colourings, in particular one which might yield a higher value of $\mu(X)$. Details on the computational part of the proof will be discussed in Section 9.4.

9.2 Defining a valid $(\mathcal{A}, \mathcal{F})$-set

We now describe how to construct a valid $(\mathcal{A}, \mathcal{F})$-set S. Let $\mathcal{R}_{\text{big}}$ be the extended region in Figure 9.4(a). For any edge-boundary pair X, the region R_X will match $\mathcal{R}_{\text{big}}$ with respect to e_X for a suitable combination of the labels "in" and "out" on the vertices of $\mathcal{R}_{\text{big}}$.

Fix any edge-boundary pair X, and suppose $\mathcal{R}_{\text{big}}$ (with some "in" and "out" labels) is the extended region matching R_X. Suppose that X belongs to the set M_m, where $m \in \{1, \ldots, 4\}$. We will construct an element $(A_{a,m}, F_{f,m}, \mathcal{H})$ in S. Let $A_{a,m}$ be the unique set of edge-boundary pairs in $\mathcal{A}$ such that $\mathcal{R}_{A_a}$ matches R_X with respect to e_X, and hence $X \in A_{a,m}$. Note that all vertices of $\mathcal{R}_{A_a}$ are contained in the region in Figure 9.2(a), which is repeated in Figure 9.4(b). The extended region $\mathcal{R}_{A_a}$ is the subregion of $\mathcal{R}_{\text{big}}$, taken such that the designated edges coincide. All vertices of $\mathcal{R}_{A_a}$ are therefore contained within the thick border in Figure 9.4(c), which illustrates how $\mathcal{R}_{A_a}$ is a subregion of $\mathcal{R}_{\text{big}}$.

For any pair of distinct colours $c, c' \in Q$, and $i \in \{1, 2, 3\}$, let $X_i(c, c')$ be the edge-boundary pair that is constructed recursively in the tree T_X, where i is the index in the clockwise ordering of the neighbours of v_X, starting with the vertex after w_X. If the ith neighbour of v_X is not in R_X, then let $a_i = 0$, otherwise let $a_i \in \{1, \ldots, 342\}$

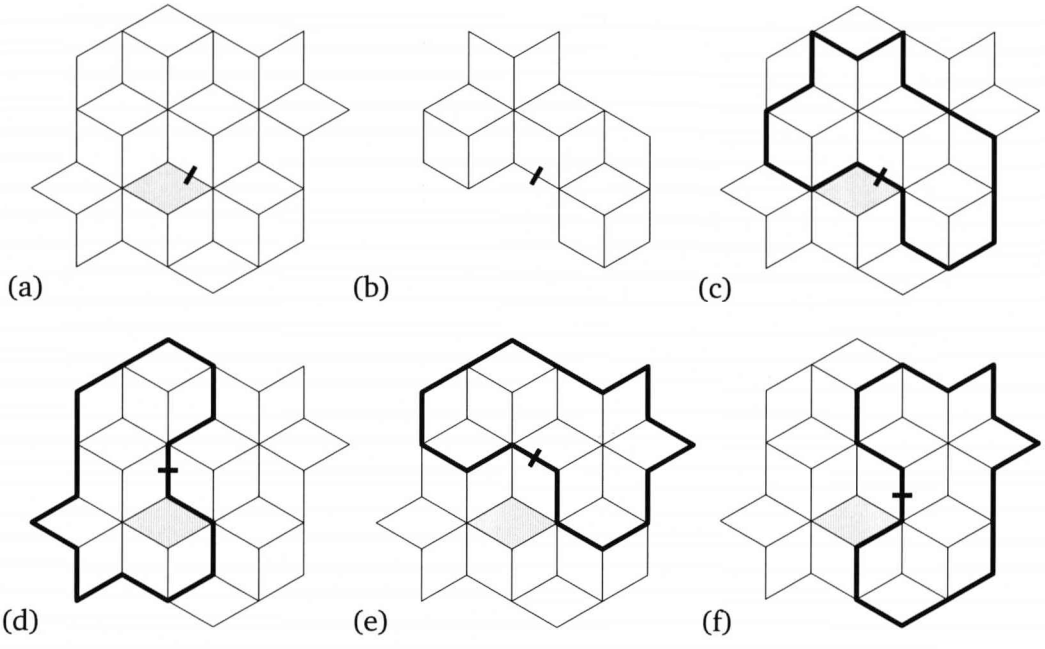


Figure 9.4. (a) The extended region $\mathcal{R}_{\text{big}}$ with all vertices labelled "in". Note that the shaded vertex is not a part of the region. (b) The extended region $\mathcal{R}_A$ repeated. (c)–(f) The extended region $\mathcal{R}_A$, outlined with a thick border, is a subregion of $\mathcal{R}_{\text{big}}$, taken such that the designated edge of $\mathcal{R}_A$ coincides with different edges of $\mathcal{R}_{\text{big}}$.

be the index such that $\mathcal{R}_{A_{a_i}}$ matches R_X with respect to the edge between v_X and its ith neighbour. That is, $\mathcal{R}_{A_{a_i}}$ is the subregion of $\mathcal{R}_{\text{big}}$ contained within the thick border in Figure 9.4(d), (e) and (f), depending on which neighbour of v_X is considered. It follows that the edge-boundary pair $X_i(c, c')$ belongs to the set A_{a_i}, provided the ith neighbour of v_X is in R_X. In order to derive which set $A_{a_i,m_i} \in \mathcal{A}$ that $X_i(c, c')$ belongs to, we must take into account the colours c and c' and the set M_m to which X belongs. We do the following case study.

Consider Figure 9.5 in which twelve cases are listed. Each case illustrates a set of edge-boundary pairs X, where, for each neighbour of v_X, shading specifies whether or not the neighbour is in R_X. Also the colours of some boundary edges are specified (c and c' in the diagram). We will see that an edge-boundary pair X belongs to at least one of these cases, depending on which neighbours of v_X are in R_X and to which set M_m X belongs. Suppose the colour of edge e_X is c in B_X and c' is B'_X, and suppose the weight $\omega_c \geq \omega_{c'}$. That is, $\mu(X) = \mu_{c,c'}(X) \geq \mu_{c',c}(X)$. Note that, since $\lambda = 0$, the probability $p_X^{\min}(c_1, c_2)$ is strictly positive only if $c_2 = c$. That is, if (C, C') is a pair of colourings of R_X drawn from $\Psi_X^{\min}$, then $C(v_X) \neq C'(v_X)$ only if $C'(v_X) = c$. If $C'(v_X) \neq c$, then we would choose $C(v_X) = C'(v_X)$ under the coupling $\Psi_X^{\min}$.

From above we had that $X \in A_{a,m}$ and $X_i(\cdot, \cdot) \in A_{a_i,m_i}$, where a and a_i are uniquely specified by the region $\mathcal{R}_{\text{big}}$. As an example, suppose that three neighbours of v_X are in R_X and suppose that $X \in M_1$ (that is, $m = 1$). Then X belongs to the case

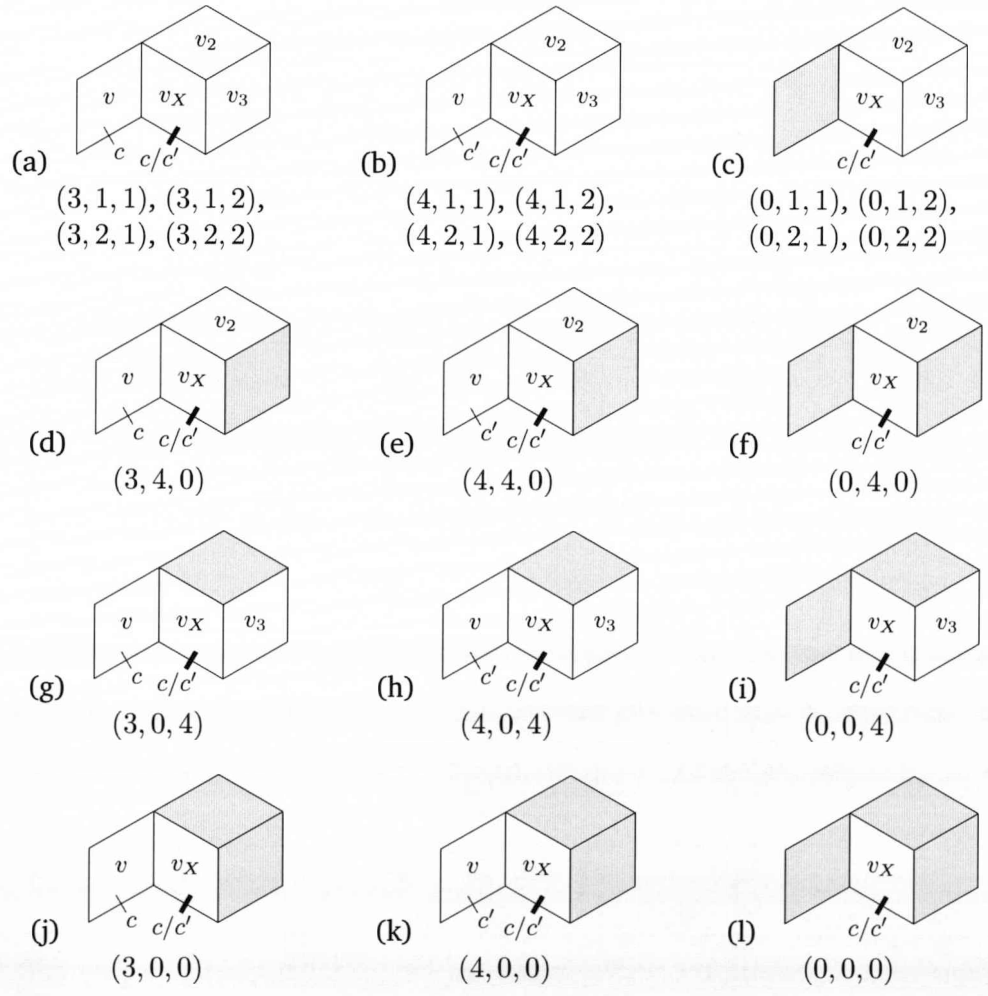


Figure 9.5. Twelve cases which cover all possible combinations of the sets $M_1, \ldots, M_4$ to which the recursively constructed edge-boundary pairs $X_1(c_1, c_2)$, $X_2(c_1, c_2)$ and $X_3(c_1, c_2)$ belong.

in Figure 9.5(a). Let $c_1, c_2 \in Q$ be two distinct colours such that $p_X^{\min}(c_1, c_2) > 0$. Then the recursively constructed edge-boundary pair $X_1(c_1, c_2)$ must belong to M_3 since the colour of $e_{X_1(c_1,c_2)}$ is c in either $B_{X_1(c_1,c_2)}$ or $B'_{X_1(c_1,c_2)}$. The other two recursively constructed edge-boundary pairs $X_2(c_1, c_2)$ and $X_3(c_1, c_2)$ will belong to M_1 or M_2. Either both belong to M_1 or both belong to M_2, or one belongs to M_1 and the other to M_2. Hence there are four possible combinations of the values of m_1, m_2 and m_3 such that $X_i(c_1, c_2) \in A_{a_i,m_i}$. Namely, (m_1, m_2, m_3) is $(3, 1, 1)$, $(3, 1, 2)$, $(3, 2, 1)$ or $(3, 2, 2)$. These are listed under the diagram in Figure 9.5(a). For each of these tuples we construct a 3-tuple that contains the sets to which $X_1(c_1, c_2), \ldots, X_3(c_1, c_2)$ could belong. These 3-tuples are $(A_{a_1,3}, A_{a_2,1}, A_{a_3,1})$, $(A_{a_1,3}, A_{a_2,1}, A_{a_3,2})$, $(A_{a_1,3}, A_{a_2,2}, A_{a_3,1})$ and $(A_{a_1,3}, A_{a_2,2}, A_{a_3,2})$, and they define the set $\mathcal{H}$ in the element $(A_{a,m}, F_{f,m}, \mathcal{H})$ in the $(\mathcal{A}, \mathcal{F})$-set S. Similarly, for all other cases in Figure 9.5, we specify possible values of m_1, m_2 and m_3. If the neighbour v_i of v_X is not in R_X then we define $m_i = 0$ and $A_{0,0} = \emptyset$. Thus, from the extended region $\mathcal{R}_{\text{big}}$ (with some "in" and "out" labels) and

the value $m \in \{1, \ldots, 4\}$, we use the case analysis in Figure 9.5 to derive the set $\mathcal{H}$ of 3-tuples, containing the sets in $\mathcal{A}$ to which the recursively constructed edge-boundary pairs $X_1(\cdot,\cdot), \ldots, X_3(\cdot,\cdot)$ could belong. It remains to derive the set $F_{f,m} \in \mathcal{F}$ in order to complete the construction of the element $(A_{a,m}, F_{f,m}, \mathcal{H})$ in S.

Let $\mathcal{F}' \subseteq \mathcal{F}$ contain the sets $F_{i,m}$ in $\mathcal{F}$, such that $X \in F_i$. That is, $F_{i,m} \in \mathcal{F}$ is in $\mathcal{F}'$ if $\mathcal{R}_{F_i}$ is a subregion of $\mathcal{R}_{\text{big}}$ such that the designated edge in $\mathcal{R}_{F_i}$ and $\mathcal{R}_{\text{big}}$ coincide, and all vertices of $\mathcal{R}_{F_i}$ are labelled "in" in $\mathcal{R}_{\text{big}}$. Let $F_{f,m}$ be the set in $\mathcal{F}'$ that minimises $\mu_{F_{i,m}}$ over all $F_{i,m} \in \mathcal{F}'$, where $\mu_{F_{i,m}}$ is the value in Lemma 9.1. From the lemma we have that $\mu(X) \leq \mu_{F_{f,m}}$.

The procedure above describes how the element $(A_{a,m}, F_{f,m}, \mathcal{H})$ in the $(\mathcal{A}, \mathcal{F})$-set S is derived from an extended region $\mathcal{R}_{\text{big}}$ (with some "in" and "out" labels) and a value $m \in \{1, \ldots, 4\}$. The procedure is implemented as a computer program, which generates the whole set S by looping through all extended regions $\mathcal{R}_{\text{big}}$ and values of m. From the construction of S it follows that S is valid. The next lemma, which is proved by computer assistance, says that S is good with respect to the values $\mu_{F_{f,m}}$ in Lemma 9.1.

Lemma 9.2. *The $(\mathcal{A}, \mathcal{F})$-set S is good with respect to $\varepsilon = 1/1000$ and the values $\mu_{F_{f,m}}$ in Lemma 9.1.*

Proof. Lemma 9.1 says that $\max_{X \in F_{f,m}} \mu(X) \leq \mu_{F_{f,m}}$ for $f \in \{1, \ldots, 4720\}$ and $m \in \{1, \ldots, 4\}$. For every $A_{a,m} \in \mathcal{A}$, let $2 \leq \alpha_{A_{a,m}} \leq 5$ be constants (that are not listed here). See Section 9.4 for more on the computational part of the proof. For each element $(A_{a,m}, F_{f,m}, \mathcal{H}) \in S$, we successfully verify for each 3-tuple $(A_{a_1,m_1}, A_{a_2,m_2}, A_{a_3,m_3}) \in \mathcal{H}$ that $\mu_{F_{f,m}}(\alpha_{A_{a_1,m_1}} + \alpha_{A_{a_2,m_2}} + \alpha_{A_{a_3,m_3}}) \leq \alpha_{A_{a,m}}(1 - \varepsilon)$. □

Strong spatial mixing now follows from Lemma 9.2. Next is the proof of Theorem 3.4. First we restate the theorem.

Theorem (3.4). *The anti-ferromagnetic Potts model on the kagome lattice with parameters $q = 5$ and $\lambda = 0$ has strong spatial mixing.*

Proof. Let $\varepsilon = 1/1000$. From Lemma 9.2 and Lemma 6.12 it follows that there is a constant $\alpha \geq 0$ such that $\Gamma^d(X) \leq \alpha(1 - \varepsilon)^d$ for every edge-boundary pair X. Strong spatial mixing now follows from Lemma 6.3. □

9.3 Rapidly mixing Glauber dynamics

Since the number of colours $q = 5$ is less than $\Delta + 2$, we cannot apply Lemma 6.3 to show that the Glauber dynamics is rapidly mixing. However, we conclude from Lemmas 6.2, 5.3, 2.23 and 5.5 that the spin system has a K-coupling cover. Thus, by

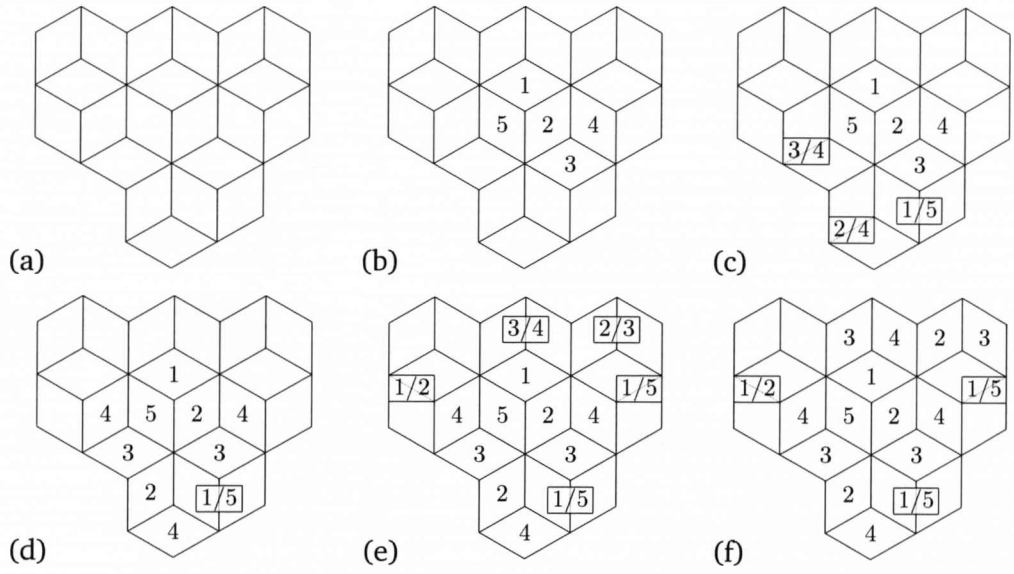


Figure 9.6. In every proper 5-colouring the region in (a) there is a vertex that has two neighbours with the same colour. (a)–(f) A step-by-step procedure that colours the region and eventually forces us to introduce a vertex that has two neighbours with the same colour.

Lemma 4.5 there is an integer $d \geq 0$ such that $\mathcal{M}_d$ is rapidly mixing on any region R and q_0-colouring $\mathcal{B}$ of ∂R. Recall from Section 2.3.1, in particular Figure 2.2, that there are regions of the kagome lattice for which the Glauber dynamics is not irreducible if $q = 5$ and the colouring of the boundary is an arbitrary q_0-colouring. However, as we will see, if we restrict the colouring of the boundary to the free boundary (the 0-colouring) then the Glauber dynamics is irreducible.

For a region R, a q-colouring σ of R and a q_0-colouring $\mathcal{B}$ of ∂R, a *single-vertex update* of a vertex $v \in R$ means that we obtain a new q-colouring σ' of R from σ by changing the colour of v to a colour in Q that is not used by the neighbours of v (under σ or $\mathcal{B}$). In order to show that the Glauber dynamics is rapidly mixing under the 0-colouring of the boundary, we will show that at most a constant number of single-vertex updates are required to go from one colouring σ_1 of the region to another colouring σ_2, where σ_1 and σ_2 differ at only a constant number of vertices. We will use this fact to provide an alternative proof of Lemma 4.8, which requires that $q \geq \Delta + 2$ when $\lambda = 0$. That is, we will show that the congestion $A(\phi)$ is bounded by a constant in the comparison between the Markov chain $\mathcal{M}_d$ and $\mathcal{M}_0$, even for $q = 5$. Once we have established this fact, we show rapid mixing by applying existing lemmas. Here follows two lemmas that we use to construct a series of single-vertex updates to go from a colouring σ_1 to σ_2.

Lemma 9.3. *Consider the region in Figure 9.6(a). In every proper 5-colouring of this region there is a vertex that has two neighbours with the same colour.*

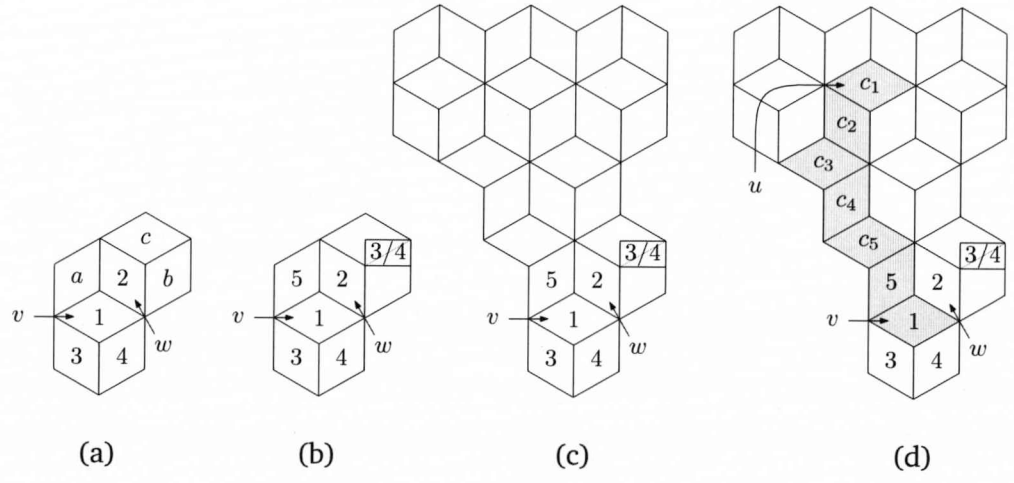
Proof. We will try to colour the region such that no vertex have two neighbours with the same colour, and we will see that it impossible to do so. First we assign colours to five of the vertices of the region, making sure that no vertex has two neighbours with the same colour. Without loss of generality we choose colours according to Figure 9.6(b). A vertex is labelled with its colour. It follows that the two vertices adjacent to the vertex coloured 5 must receive colour 3 and 4. Otherwise we would have a vertex that has two neighbours with the same colour. Similarly, the vertices adjacent to the vertex coloured 3 must receive colour 1 and 5, and therefore the two bottom left vertices are forced to receive colour 2 and 4. Figure 9.6(c) illustrates this fact, where a square contains the two colours that have to be assigned to the two vertices it is overlapping. From the two leftmost squares we see that the colour 4 has to be assigned to the vertices that are as far apart as possible, forcing us to choose the colouring in Figure 9.6(d). Figure 9.6(e) illustrates how we are forced to assign colours to other vertices of the region, and Figure 9.6(f) shows how the four rightmost vertices at the top must be coloured. To finish the proof we note that it is now impossible to colour the two leftmost vertices at the top without introducing a vertex such that two of its neighbours receive the same colour. □

Lemma 9.4. *Let R be a region of the kagome lattice and let $\mathcal{B}$ be the 0-colouring of the boundary ∂R. Suppose that $q = 5$ and let σ_1 and σ_2 be any two proper q-colourings of R that differ on m vertices. We can go from σ_1 to σ_2 by applying a series of $O(m)$ single-vertex updates.*

Proof. Let $v \in R$ be a vertex on which σ_1 and σ_2 differ. We will show how to recolour v to the colour it has in σ_2 by doing at most a constant number of single-vertex updates. A vertex in R that has the same colour in both σ_1 and σ_2 will not change colour after v has been updated. First we analyse situations where no boundary vertices in ∂R are involved. We note at the end of the proof that if boundary vertices are present, then it only makes it easier to recolour v. That is, assume for now that all vertices we consider belong to the region R. The proof goes through a series of cases.

If possible, simply recolour v to the colour it has in σ_2. If this is not possible then there must be one or two neighbours of v that have colour $\sigma_2(v)$ in σ_1. It cannot be more than two such neighbours since σ_1 is a proper colouring.

Without loss of generality, assume that $\sigma_1(v) = 1$ and $\sigma_2(v) = 2$. If v has two neighbours with colour 2 in σ_1 then we will first recolour one of these two neighbours to some other colour than 2. Let w be the neighbour of v with colour 2 that we are going to recolour. Note that $\sigma_2(w) \neq 2$ since σ_2 is a proper colouring. If possible, recolour w to some other colour than 2. If this is not possible then w is "locked" and must have three neighbours coloured 3, 4 and 5, respectively. In this case, first recolour v (which is possible since v has two neighbours with colour 2) and then recolour w to



(a) (b) (c) (d)

Figure 9.7. The colours 1 and 2 are going to swap place. Lemma 9.4 guarantees that this can be done with a constant number of single-vertex updates.

colour 1. Now only one neighbour of v has colour 2. We deal with this case next.

Without loss of generality, assume that $\sigma_1(v) = 1$ and $\sigma_2(v) = 2$, and exactly one neighbour w of v has colour 2 in σ_1. Note that $\sigma_2(w) \neq 2$ since σ_2 is a proper colouring. If possible, recolour w to something else than 2 and then recolour v to 2. If this is not possible then w is "locked" and must have four neighbours (including v) with colours 1, 3, 4 and 5, respectively, in σ_1. Without loss of generality, consider the region in Figure 9.7(a), which is a subregion of R. Call this region R'. The vertices of R' are labelled with their colours in σ_1. The vertex with colour 1 is v and the vertex with colour 2 is w. We assume without loss of generality that the two neighbours of v that are below v are the two neighbours with colour 3 and 4 in σ_1.

Three of the vertices in R' are given the colours a, b and c, which are to be determined. Since w is "locked", the colours a, b and c is any permutation of the colours 3, 4 and 5. If a is 3 or 4 then we recolour v to 5 and then recolour w to 1, and then recolour v to 2. If this is not the case then a must be 5, and hence the colours b and c are 3 and 4 in any order. Figure 9.7(b) illustrates this. We now analyse this case.

We will use Lemma 9.3 to show that we can recolour v to 2 without changing the colour of any other vertex except w (which will be recoloured to 1). Consider Figure 9.7(c) which illustrates the region R' extended with vertices in R. The vertices we extend R' with correspond to the region that we used in Lemma 9.3. From Lemma 9.3 we know that there must be at least one vertex u among the vertices we extend R' with such that u has two neighbours with the same colour. Let P be a shortest path from v to u such that the path goes from v to the neighbour above that has colour 5 and then is entirely inside the region we added to R'. Figure 9.7(d) illustrates an example of such a path. The path is shaded in the figure. Suppose that the vertex u is chosen such that all vertices on the path P (except from u itself) are "locked" (have four neighbours of

different colours). Note that if the vertex coloured 5 above v does not have four neighbours of different colours then we let u be this vertex and hence the path P consists only of the two vertices u and v.

Suppose that the path P contains k vertices. Let $c_1, \ldots, c_k$ be the colours in σ_1 of the vertices from u to v along the path. That is, $\sigma_1(u) = c_1$, $c_{k-1} = 5$ and $c_k = 1$. Since u has two neighbours with the same colour, we recolour u from c_1 to another colour c_1'. Now the vertex after u on P has two neighbours with the same colour (namely c_1'), since all its neighbours had different colours before recolouring u. We recolour this vertex from c_2 to c_2'. We continue this recolouring procedure along the path P all the way to vertex v, which will be recoloured to 3. Note that the vertex above v which had previously colour 5 now must have colour 3 or 4. We can now recolour w to 1 and then recolour v to 2. It remains to recolour the vertices on the path back to their original colours in σ_1. We do this by reversing the recolouring procedure, starting with the vertex above v, which is recoloured back to 5. When u is recoloured back to c_1 we are done.

We have now shown how a constant number of single-vertex updates are applied in order to recolour a vertex v to the colour it has in σ_2 without changing the colour on vertices that have the same colour in σ_1 and σ_2.

We note that if any vertices involved in the recolouring procedure of v are boundary vertices then this will only make it easier. Note from the statement of the lemma that we assume that a boundary vertex has colour 0. As we have seen, the tricky situations arise when a vertex is "locked" with four neighbours of different colours (excluding colour 0). Such a vertex is tricky because we cannot just change its colour to another colour in Q. A vertex that is adjacent to a boundary vertex can never be "locked" since there is always at least one colour in Q that it can be recoloured to. Thus, although the part of the proof above assumes that all vertices are in R, we note that the presence of boundary vertices only makes the recolouring procedure easier. Of course, depending on which vertex v we are going to recolour, and which neighbour w is "locked", the path P might go in a direction that is different from the one in Figure 9.7(d). However, the same technique is applied in order successfully to recolour v.

Finally, in order to go from σ_1 to σ_2, we recolour each vertex v at which σ_1 and σ_2 differ. For each such vertex it takes only a constant number of single-vertex updates to do so. Since σ_1 and σ_2 differ only at m vertices, the total number or updates is $O(m)$. □

We are now ready to prove the following lemma, which replaces Lemma 4.8 in the proof of rapidly mixing Glauber dynamics for five colours on the kagome lattice. Recall the definition of the congestion $A(\phi)$ in Section 4.4.

Lemma 9.5. *Consider any region R of the kagome lattice and let the colouring $\mathcal{B}$ of ∂R be the 0-colouring. If $q = 5$ and $\lambda = 0$, then there is a flow ϕ such that the congestion $A(\phi) \in O(1)$.*

Proof. The proof is identical to the proof of Lemma 4.8, only with the difference that the path $\gamma_{\sigma_1,\sigma_2} \in \mathcal{P}_{\sigma_1,\sigma_2}$ from a colouring σ_1 of the region to another colouring σ_2 is constructed differently. Here we define the path $\gamma_{\sigma_1,\sigma_2}$ by applying Lemma 9.4. We update vertices in the order specified by $\prec$. It follows that the path length $|\gamma_{\sigma_1,\sigma_2}|$ is upper-bounded by a constant since σ_1 and σ_2 differ only on vertices inside a ball of fixed radius d. For each vertex v that is updated, we do a constant number of recolourings of vertices that are within constant distance from v. With these modifications and observations, the proof of Lemma 4.8 also proves this lemma. □

We are now able to show that the Glauber dynamics is rapidly mixing for five colours on the kagome lattice. Next is the proof of Theorem 3.5. The theorem is restated below.

Theorem (3.5). *For any region R of the kagome lattice, $q = 5$ and $\lambda = 0$, the Glauber dynamics is rapidly mixing for the 0-colouring $\mathcal{B}$ of ∂R. The mixing time $\tau(\delta) \in O(n^2 + n \log \frac{1}{\delta})$, where n is the number of vertices in R.*

Proof. From Lemma 9.2 and Lemma 6.12 it follows that there is a constant $\alpha \geq 0$ such that $\Gamma^d(X) \leq \alpha(1-\varepsilon)^d$ for every edge-boundary pair X, where $\varepsilon = 1/1000$. From Lemmas 6.2, 5.3, 2.23 and 5.5 it follows that the system has a K-coupling cover. From Lemma 4.5 it follows that $\mathcal{M}_d$ is rapidly mixing on any region R for some integer $d \geq 0$. From Lemma 9.5 we have that the congestion $A(\phi) \in O(1)$ if the colouring $\mathcal{B}$ of ∂R is the 0-colouring, and hence Lemma 4.7 implies that the Glauber dynamics $\mathcal{M}_0$ is rapidly mixing with the stated mixing time. □

9.4 The computational part

The computational part of the proof of exponential decay consists of two tasks: computing the values $\mu_{F_{f,m}}$ for $f = 1, \ldots, 4720$ and $m = 1, \ldots, 4$ in Lemma 9.1, and constructing the $(\mathcal{A}, \mathcal{F})$-set S and verifying that it is good with respect to the values $\mu_{F_{f,m}}$. Like for the triangular lattice (see Section 8.2.1), the latter task is rather straightforward. We have implemented a program in standard C that constructs the set S. The program outputs a file which represents the linear program that contains all inequalities that have to be satisfied in order to make S good with respect to the values $\mu_{F_{f,m}}$. The linear program contains the 684 unknown variables $\alpha_{A_{a,m}}$, where $A_{a,m} \in \mathcal{A}$. We use the GNU Linear Programming Kit to find a solution, which is found with the constant ε set to $1/1000$ and each variable $\alpha_{A_{a,m}}$ being in the interval $[2, 5]$. Before actually feeding

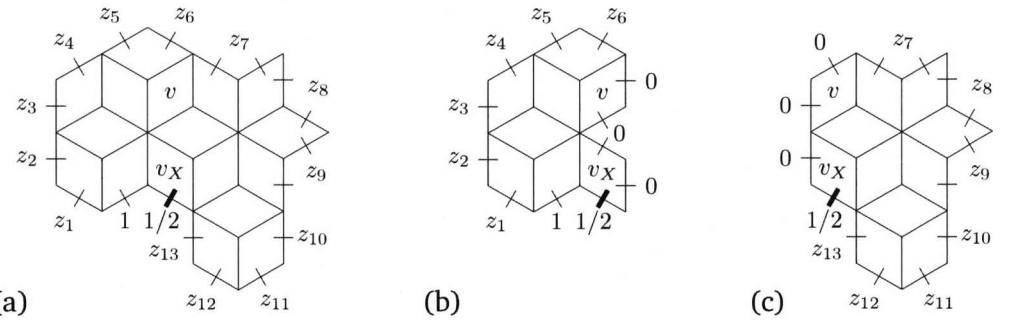


Figure 9.8. The region R_X of the edge-boundary pair X in (a) is split into two: (b) and (c). The split is along the vertices labelled v and v_X.

the linear program to the GNU Linear Programming Kit, we filter out inequalities that are redundant. We have written a C-program that goes through the linear program and checks if there is a pair of inequalities, say $x < y_1$ and $x < y_2$, such that $y_1 \leq y_2$. If there is such a pair of inequalities then it is unnecessary to keep the inequality $x < y_2$. This inequality will be removed from the linear program. After removing redundant inequalities, we have a linear program that contains around 160,000 inequalities. For details on the implementation, see webpage `http://www.csc.liv.ac.uk/~markus/phdthesis/`

9.4.1 Computing $\mu_{f,m}$

We face a serious problem when computing the 9440 values $\mu_{F_{f,m}}$. We must be able to compute a single value in rather short time, otherwise the total running time for all values will be too long. Using the same approach as for the triangular lattice in Section 8.2 is not feasible here. Instead we use a technique that is explained with the following example.

Consider an edge-boundary pair $X \in F_{f,m}$. Suppose $\mathcal{R}_{F_f}$ is the extended region in Figure 9.8(a), and suppose that $m = 1$ (meaning $X \in M_1$). In order to compute $\mu_{F_{f,m}}$, we will maximise $\mu(X)$ over pairs of colourings (B_X, B'_X) of the boundary. We assume without loss of generality that $B_X(e_X) = 1$ and $B'_X(e_X) = 2$. Since $X \in M_1$, we want to maximise $\mu_{1,2}(X)$, where the boundary edge between ω_X and the left neighbour of v_X is given colour 1. Figure 9.8(a) illustrates the edge-boundary pair X where boundary edges are labelled with their colour. Thus, in order to maximise $\mu_{1,2}(X)$, we loop through all combinations of the colours $z_1, \ldots, z_{13}$ and compute $\mu_{1,2}(X)$ for each such combination. We make the following observation. If, say $z_1, \ldots, z_6$ are all colour 1, then $\mu_{1,2}(X)$ will most likely not be maximised, regardless of the remaining colours $z_7, \ldots, z_{13}$. That is, setting all the colours $z_1, \ldots, z_6$ to 1 is a "stupid" choice if we want to maximise $\mu_{1,2}(X)$. The idea we explain next captures this observation by filtering out certain "stupid" colourings of the boundary, and thereby speeding up the process of finding the highest value $\mu_{1,2}(X)$.

We "split" the extended region $\mathcal{R}_{F_f}$ into two extended regions $\mathcal{R}^{\text{left}}$ and $\mathcal{R}^{\text{right}}$. Figure 9.8(b) and (c) illustrate $\mathcal{R}^{\text{left}}$ and $\mathcal{R}^{\text{right}}$, respectively, which are obtained by splitting the region in Figure 9.8(a). The two regions share the vertices in the split. In this case they are the vertices v_X and v, both labelled in the figure. Let X^{left} and X^{right} be the two edge-boundary pairs with region $\mathcal{R}^{\text{left}}$ and $\mathcal{R}^{\text{right}}$, respectively, and for which the boundary edges receive the same colouring as in X. Boundary edges that are introduced from the split are given colour 0. Figure 9.8(b) and (c) illustrate X^{left} and X^{right}.

Recall from Section 6.2 that when $\lambda = 0$, the weight ω_c of the edge-boundary pair X is exactly the number of proper colourings σ of R_X that agree with the boundary (excluding edge e_X) such that the colour $\sigma(v_X) = c$. We now define the weight $\omega_{c,c'}$ as the number of proper colourings σ of R_X that agree with the colouring of the boundary (excluding edge e_X) such that $\sigma(v_X) = c$ and $\sigma(v) = c'$, where v is the second vertex in the split. That is,

$$\omega_c = \sum_{c'=1}^{q} \omega_{c,c'}.$$

The weights $\omega_{c,c'}^{\text{left}}$ and $\omega_{c,c'}^{\text{right}}$ are defined similarly for X^{left} and X^{right}. It now follows that

$$\omega_{c,c'} = \omega_{c,c'}^{\text{left}} \omega_{c,c'}^{\text{right}},$$

and hence

$$\omega_c = \sum_{c'=1}^{q} \omega_{c,c'}^{\text{left}} \omega_{c,c'}^{\text{right}}.$$

With $q = 5$ colours, we have

$$\begin{aligned}
\mu_{1,2}(X) &= \frac{\omega_1}{\sum_{i\in\{1,3,4,5\}} \omega_i} = \frac{\sum_{j=1}^{5} \omega_{1,j}^{\text{left}} \omega_{1,j}^{\text{right}}}{\sum_{i\in\{1,3,4,5\}} \sum_{k=1}^{5} \omega_{i,k}^{\text{left}} \omega_{i,k}^{\text{right}}} \\
&= \sum_{j=1}^{5} \frac{\omega_{1,j}^{\text{left}} \omega_{1,j}^{\text{right}}}{\sum_{i\in\{1,3,4,5\}} \sum_{k=1}^{5} \omega_{i,k}^{\text{left}} \omega_{i,k}^{\text{right}}} = \sum_{j=1}^{5} \frac{1}{\sum_{i\in\{1,3,4,5\}} \sum_{k=1}^{5} \frac{\omega_{i,k}^{\text{left}} \omega_{i,k}^{\text{right}}}{\omega_{1,j}^{\text{left}} \omega_{1,j}^{\text{right}}}} \\
&= \sum_{j=1}^{5} \frac{1}{\sum_{i\in\{1,3,4,5\}} \sum_{k=1}^{5} \left(\frac{\omega_{i,k}^{\text{left}}}{\omega_{1,j}^{\text{left}}} \cdot \frac{\omega_{i,k}^{\text{right}}}{\omega_{1,j}^{\text{right}}} \right)}. \qquad (9.1)
\end{aligned}$$

Let B^{left} be a choice of the colours $z_1, \ldots, z_6$, and let B^{right} be a choice of the colours $z_7, \ldots, z_{13}$. Thus, B^{left} defines the colouring of the boundary of X^{left} and hence the weights $\omega_{c,c'}^{\text{left}}$. Similarly, B^{right} defines the colouring of the boundary of X^{right} and the weights $\omega_{c,c'}^{\text{right}}$. In order to maximise $\mu_{1,2}(X)$ over all colourings of the boundary, we consider all combinations of colours B^{left} and B^{right}, and use Equation (9.1). Now,

let $B^{\text{left-1}}$ and $B^{\text{left-2}}$ be two different choices of the colours $z_1, \ldots, z_6$, such that

$$\frac{\omega_{i,k}^{\text{left-1}}}{\omega_{1,j}^{\text{left-1}}} \leq \frac{\omega_{i,k}^{\text{left-2}}}{\omega_{1,j}^{\text{left-2}}}$$

for all $i \in \{1,3,4,5\}$, $j \in \{1,\ldots,5\}$ and $k \in \{1,\ldots,5\}$. Then we have from Equation (9.1) that $\mu_{1,2}(X)$ can only get smaller if we take $B^{\text{left}} = B^{\text{left-2}}$ instead of $B^{\text{left}} = B^{\text{left-1}}$. In other words, there is no point considering the colours defined by $B^{\text{left-2}}$ when maximising $\mu_{1,2}(X)$. This observation suggests that we loop through all colours B^{left}, and perform pairwise comparisons, and keep only those that cannot be ruled out like $B^{\text{left-2}}$ above. This gives us a set of colours $z_1, \ldots, z_6$ that turns out to be much smaller than the set containing all combinations of colours. Similarly, we obtain a set of colours $z_7, \ldots, z_{13}$ by looping through colours B^{right} and performing pairwise comparisons. We combine the two sets as in Equation (9.1) in order to find the maximum value $\mu_{1,2}(X)$. As described in previous chapters, the weights ω_c (and here $\omega_{c,c'}$) are obtained by using a dynamic programming approach.

The technique of splitting regions and filtering out colourings that are guaranteed not to maximise $\mu_{1,2}(X)$ turns out to have a huge impact on the running time of the program. We have tried to compute μ-values in the same manner as for the triangular lattice. That is, by looping through colourings of the whole boundary of a region, and for each boundary colouring compute the μ-value. We wrote a C-program that did this for some of the larger regions. While running the program we monitored the progress of which boundary colouring was being processed. We could immediately tell that the speed at which the program was working through boundary colourings would result in a running time of several months (for the larger regions). When implementing the technique of splitting the regions we managed to finish all computations is about two days. We run the program on a fairly powerful home-PC as of year 2006. We left the computer running over night without disturbing it. It should be mentioned that a downside with the splitting-technique is the amount of time it takes to implement the program since the program is rather involved. A vast amount of time has been spent implementing and making sure the program is correct and reliable.

Chapter 10

Systematic scan on the grid

In this chapter we prove Theorem 3.6. We will make use of a recent result by Pedersen [Ped07] to prove rapid mixing of systematic scan. We will use a heuristic to mechanically construct sufficiently good couplings of uniform distributions on proper colourings of a 2×2-block. We will hence use a heuristic-based computation in order to establish a rigorous result about the mixing time of a systematic scan Markov chain.

This chapter is organised as follows. In Section 10.1 we provide some definitions and notations that are used specifically in the context of systematic scan. In Section 10.2 we describe the method we use in order to show that the systematic scan Markov chain is rapidly mixing. Here we present the proof of Theorem 3.6. In Section 10.3 we explain how to construct the couplings that were used in this proof. Finally, in Section 10.4, we consider the possibility of proving rapid mixing of systematic scan for 6-colourings of the grid by increasing the size of the blocks. We show that the proof technique we employ does not imply rapid mixing of systematic scan for 6-colourings of the grid when using 2×2-, 2×3- and 3×3-blocks.

10.1 Preliminaries

Let R be a torus (see the end of Section 2.4.1 on page 25). Note that $\partial R = \emptyset$. By working on the torus we avoid treating several technicalities regarding the vertices on the boundary of an arbitrary region of $\mathbb{Z}^2$. This lets us present the proof of rapid mixing in a more "clean" way. In this chapter, let $\Omega^+ = \Omega_R^+$ denote the set of all q-colourings of R and $\Omega = \Omega_R$ denote the set of proper q-colourings of R. Let π be the uniform distribution on Ω. For each vertex $u \in R$, let $\Omega^{+2}(u)$ denote the set of pairs $(\sigma_1, \sigma_2) \in \Omega^+ \times \Omega^+$ of colourings that only differ on the colour assigned to vertex u. That is, $(\sigma_1, \sigma_2) \in \Omega^{+2}(u)$ if and only if $\sigma_1(u) \neq \sigma_2(u)$ and $\sigma_1(v) = \sigma_2(v)$ on all vertices $v \in R \backslash \{u\}$. Recall that an $n \times m$-block is a rectangular region of the square lattice that has height n vertices and width m vertices.

Recall Definition 2.16 of a general systematic scan Markov chain $\mathcal{M}_{\text{scan}}$. The systematic scan Markov chain we use here (denoted $\mathcal{M}_{\text{scan}}$) is specified as follows. Each subset $\Theta \in \{\Theta_1, \ldots, \Theta_m\}$ of the vertices of R is a 2×2-block. We will choose the number m of blocks to be as small as possible. In Section 10.4 we will consider systematic scan Markov chains that are defined with bigger block sizes. For each block Θ_k, $k \in \{1, \ldots, m\}$, and any two colourings $\sigma_1, \sigma_2 \in \Omega^+$, we write "$\sigma_1 = \sigma_2$ on Θ_k" if $\sigma_1(v) = \sigma_2(v)$ for each $v \in \Theta_k$. Similarly we write "$\sigma_1 = \sigma_2$ off Θ_k" if $\sigma_1(v) = \sigma_2(v)$ for each $v \in R \setminus \Theta_k$. For a block Θ_k and a colouring $\sigma \in \Omega^+$, let $\Omega_k^\sigma \subseteq \Omega^+$ be the set of colourings σ' such that $\sigma' = \sigma$ off Θ_k and $\sigma'(\Theta_k)$ is a proper colouring that agrees with the boundary $\sigma(\partial\Theta_k)$. Let π_k^σ be the uniform distribution on Ω_k^σ.

10.2 Bounding the mixing time of systematic scan

This section contains the proof of Theorem 3.6 although the proof of a crucial lemma, which requires computer assistance, is deferred to Section 10.3. We will bound the mixing time of $\mathcal{M}_{\text{scan}}$ by bounding the influence on a vertex, a parameter which we denote by α. For any two colourings $\sigma_1, \sigma_2 \in \Omega^+$, let $\Psi_k(\sigma_1, \sigma_2)$ be a coupling of the distributions $\pi_k^{\sigma_1}$ and $\pi_k^{\sigma_2}$. We let $p_v(\Psi_k(\sigma_1, \sigma_2))$ denote the probability that a vertex $v \in \Theta_k$ is assigned different colours in a pair of colourings drawn from the coupling $\Psi_k(\sigma_1, \sigma_2)$. We then let

$$\rho_{u,v}^k = \max_{(\sigma_1,\sigma_2)\in\Omega^{+2}(u)} p_v(\Psi_k(\sigma_1, \sigma_2))$$

be the influence of u on v under Θ_k, where $u \in R$ and $v \in \Theta_k$. The parameter α, denoting the influence on any vertex, is defined as

$$\alpha = \max_{k\in\{1,\ldots,m\}} \max_{v\in\Theta_k} \sum_{u\in R} \rho_{u,v}^k.$$

If α is sufficiently small then Theorem 2 from Pedersen [Ped07] implies that any systematic scan Markov chain, whose transition matrices for updating each block satisfy two simple properties, mixes in $O(\log n)$ scans, where n is the number of vertices in R. One scan takes $\sum_{k=1}^m |\Theta_k|$ vertex updates and is therefore of order $O(n)$. Note that Pedersen [Ped07] actually defines α with a weight associated with each vertex, however, as we will not use weights in our proof we have omitted them from the above account. For completeness we restate Theorem 2 from [Ped07].

Theorem 10.1 (Theorem 2, Pedersen [Ped07]). *If $\alpha < 1$ then the mixing time of $\mathcal{M}_{\text{scan}}$ is*

$$\tau_{\mathcal{M}_{\text{scan}}}(\delta) \leq \frac{\log(n/\delta)}{1-\alpha}.$$

Let P_k be the $|\Omega| \times |\Omega|$-transition matrix under a heat-bath update with respect to Θ_k, and let $P_k(\sigma_1, \sigma_2)$ denote the probability of making a transition from colouring σ_1 to σ_2. In order for Theorem 10.1 to apply, the following two properties must hold.

(1) If $P_k(\sigma_1, \sigma_2) > 0$ then $\sigma_1 = \sigma_2$ off Θ_k, and

(2) $\pi P_k = \pi$.

From Definition 2.13 of a heat-bath update and the proof of Lemma 2.17 (saying that π, the uniform distribution on Ω, is the stationary distribution of $\mathcal{M}_{\text{scan}}$) we have that both properties are satisfied. Hence we are able to use Theorem 10.1 to bound the mixing time of $\mathcal{M}_{\text{scan}}$.

So, in order to upper-bound α we are required to upper-bound the probability of a discrepancy at each vertex $v \in \Theta_k$ under a coupling $\Psi_k(\sigma_1, \sigma_2)$ for any pair of colourings $(\sigma_1, \sigma_2) \in \Omega^{+2}(u)$ that only differ at the colour of vertex $u \in R$. Our main task is hence to specify the coupling $\Psi_k(\sigma_1, \sigma_2)$ of $\pi_k^{\sigma_1}$ and $\pi_k^{\sigma_2}$ such that we have $\alpha < 1$.

Consider any block Θ_k and any pair of colourings $(\sigma_1, \sigma_2) \in \Omega^{+2}(u)$ that differ only on the colour assigned to some vertex $u \in R$. Clearly the distribution on colourings of Θ_k that is induced by $\pi_k^{\sigma_1}$ only depends on the boundary colouring $\sigma_1(\partial\Theta_k)$. Similarly, the distribution on colourings of Θ_k induced by $\pi_k^{\sigma_2}$ depends only on $\sigma_2(\partial\Theta_k)$. If $u \notin \partial\Theta_k$ then the distributions on the colourings of Θ_k, induced by $\pi_k^{\sigma_1}$ and $\pi_k^{\sigma_2}$, respectively, are the same and we let $\Psi_k(\sigma_1, \sigma_2)$ be the coupling in which any pair of colourings drawn from $\Psi_k(\sigma_1, \sigma_2)$ agree on Θ_k. That is, if the pair (σ_1', σ_2') of colourings are drawn from $\Psi_k(\sigma_1, \sigma_2)$ then $\sigma_1' = \sigma_1$ off Θ_k, $\sigma_2' = \sigma_2$ off Θ_k and $\sigma_1' = \sigma_2'$ on Θ_k. This gives $\rho_{u,v}^k = 0$ for every $u \notin \partial\Theta_k$ and $v \in \Theta_k$.

We now need to construct $\Psi_k(\sigma_1, \sigma_2)$ for the case when $u \in \partial\Theta_k$. If the values $\rho_{u,v}^k$ are too big then the parameter α will be too big (that is greater than one) and we cannot make use of Theorem 10.1. Constructing $\Psi_k(\sigma_1, \sigma_2)$ by hand such that $p_v(\Psi_k(\sigma_1, \sigma_2))$ is sufficiently small is a difficult task. It is, however, possible to use computer assistance in order to construct a coupling that is good enough for our purposes. We need to construct a specific coupling for each pair of colourings σ_1 and σ_2 that differ only at a single vertex $u \in \partial\Theta_k$. In order to do this we will make use of the following lemma, which is proved in Section 10.3.

Lemma 10.2. *Let $v_1, \ldots, v_4$ be the four vertices in a 2×2-block and let $z_1, \ldots, z_8$ be the boundary vertices of the block and let the labelling be as in Figure 10.1. Let Z and Z' be any two 7-colourings of the boundary vertices such that Z and Z' are identical on each vertex except on z_1. Let π_Z and $\pi_{Z'}$ be the uniform distributions on proper 7-colourings of the block that agree with Z and Z', respectively. For $i = 1, \ldots, 4$, let $p_{v_i}(\Psi)$ denote the probability that the colour of vertex v_i differs in a pair of colourings drawn from a*

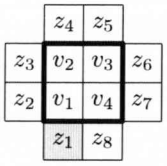


Figure 10.1. General labelling of the vertices in a 2×2-block Θ_k and on the boundary of the block.

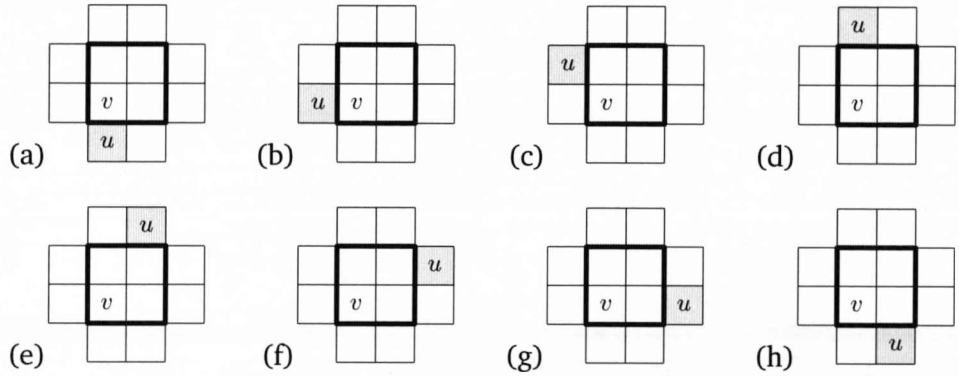


Figure 10.2. A 2×2-block Θ_k showing all eight positions of a vertex $u \in \partial\Theta_k$ on the boundary of the block in relation to a vertex $v \in \Theta_k$ in the block.

coupling Ψ of π_Z and $\pi_{Z'}$. Then there exists a coupling Ψ such that $p_{v_1}(\Psi) < 0.283$, $p_{v_2}(\Psi) < 0.079$, $p_{v_3}(\Psi) < 0.051$ and $p_{v_4}(\Psi) < 0.079$.

Thus, if $u \in \partial\Theta_k$ we let $\Psi_k(\sigma_1, \sigma_2)$ be the coupling of $\pi_k^{\sigma_1}$ and $\pi_k^{\sigma_2}$ that draws the colouring of Θ_k from the coupling Ψ in Lemma 10.2, where Z is the boundary colouring obtained from $\sigma_1(\partial\Theta_k)$ and Z' is obtained from $\sigma_2(\partial\Theta_k)$, and leaves the colour of the remaining vertices $R\backslash\Theta_k$ unchanged. That is, if the pair (σ_1', σ_2') of colourings are drawn from $\Psi_k(\sigma_1, \sigma_2)$ then $\sigma_1' = \sigma_1$ off Θ_k, $\sigma_2' = \sigma_2$ off Θ_k and the colourings of Θ_k in σ_1' and σ_2' are drawn from the coupling Ψ in Lemma 10.2. Note that due to the symmetry of the 2×2-block, with respect to rotation and mirroring, we can always label the vertices of Θ_k and $\partial\Theta_k$ such that label z_1 in Figure 10.1 represents the discrepancy vertex u on the boundary. Therefore Lemma 10.2 will provide us with upper bounds on the parameters $\rho_{u,v}^k$. We summarise the values $\rho_{u,v}^k$ in the following corollary of Lemma 10.2. Note that due to the symmetry of the block, we assume that vertex $v \in \Theta_k$ in the corollary is located in the bottom left corner, as Figure 10.2 shows.

Corollary 10.3. *Let Θ_k be any 2×2-block, let $v \in \Theta_k$ be any vertex in the block and let $u \in \partial\Theta_k$ be a vertex on the boundary of the block. Then*

$$\rho^k_{u,v} = \max_{(\sigma_1,\sigma_2)\in\Omega^{+2}(u)} p_v(\Psi_k(\sigma_1,\sigma_2)) < \begin{cases} 0.283, & \text{if } u \text{ and } v \text{ as in Figure 10.2(a) or (b),} \\ 0.079, & \text{if } u \text{ and } v \text{ as in Figure 10.2(c) or (h),} \\ 0.051, & \text{if } u \text{ and } v \text{ as in Figure 10.2(e) or (f),} \\ 0.079, & \text{if } u \text{ and } v \text{ as in Figure 10.2(d) or (g).} \end{cases}$$

If $u \notin \partial\Theta_k$ is not on the boundary of the block then $\rho^k_{u,v} = 0$.

We now use Corollary 10.3 to prove Theorem 3.6. The theorem is restated below.

Theorem (3.6). *Let R be a rectangular region of $\mathbb{Z}^2$ that is folded into a torus (hence $\partial R = \emptyset$). For $q = 7$ and $\lambda = 0$, the systematic scan Markov chain $\mathcal{M}_{\text{scan}}$ is rapidly mixing with 2×2-blocks. More precisely, the mixing time $\tau_{\mathcal{M}_{\text{scan}}}(\delta) \in O(\log \frac{n}{\delta})$, where n is the number of vertices in R.*

Proof. For every block Θ_k and $v \in \Theta_k$, let $\alpha_{k,v} = \sum_{u\in R} \rho^k_{u,v}$ be the influence on v under Θ_k. We need $\alpha_{k,v}$ to be strictly upper-bounded by one for each block Θ_k and vertex $v \in \Theta_k$ in order to ensure that $\alpha = \max_k \max_{v\in\Theta_k} \alpha_{k,v}$ is less than one. Fix any block Θ_k and any vertex $v \in \Theta_k$. A vertex $u \in \partial\Theta_k$ on the boundary of the block can occupy eight different positions on the boundary in relation to v, as showed in Figure 10.2(a)–(h). Thus, using the bounds from Corollary 10.3, we have

$$\alpha_{k,v} = \sum_{u\in R} \rho^k_{u,v} < 2(0.283 + 0.079 + 0.051 + 0.079) = 0.984,$$

which means that $\alpha = 0.984 < 1$. We obtain the stated bound on the mixing time of $\mathcal{M}_{\text{scan}}$ by applying Theorem 10.1. □

We have yet to establish a proof of Lemma 10.2, which we do in the next section. Our method of proof is computer assisted and we will be focusing on minimising the probability of assigning different colours to vertex v_1 in the constructed couplings. Our approach is similar to the one Achlioptas et al. [AMMV04] take, however we do not have the option of constructing an "optimal" coupling for each given pair of colourings of the boundary since each of the quantities $\rho^k_{u,v}$ maximises separately over the pair of boundary colourings. However, we get the mixing of systematic scan and not just of random-update dynamics.

10.3 Constructing the coupling by machine

In order to prove Lemma 10.2 we will construct a coupling Ψ of π_Z and $\pi_{Z'}$ for all pairs of boundary colourings Z and Z' that differ only at vertex z_1. For each coupling

constructed, we verify that the probabilities $p_{v_i}(\Psi)$, $i = 1, \ldots, 4$, are within the bounds of the lemma. The method is well suited to be carried out with the help of a computer and we have implemented a program in C to do so. Before stating the proof of Lemma 10.2 we will discuss how a coupling can be represented by an edge-weighted complete bipartite graph. We make use of this representation of Ψ in the proof of the lemma.

10.3.1 Representing a coupling as a bipartite graph

Let S be a set of objects and let W be a set of $|S|$ pairs (s, w_s) such that $s \in S$ and $w_s \geq 0$ is a non-negative value representing the weight of s. Each element $s \in S$ is contained in exactly one of the pairs in W. If the value w_s is an integer (which it is in our case) it can be regarded as the multiplicity of s in a multiset. The set W is referred to as a *weighted set of* S. Let $\pi_{S,W}$ be the distribution on S such that the probability of s is proportional to w_s, where (s, w_s) is a pair in W. More precisely, the probability of s in $\pi_{S,W}$ is $\pi_{S,W}(s) = w_s / \sum_{(t,w_t)\in W} w_t$.

The reason for introducing the notion of a weighted set is that it can be used when specifying a coupling of two distributions. Let S be a set and let W and W' be two weighted sets of S such that the sum of the weights in W equals the sum of the weights in W'. Let w_{tot} denote this sum. That is,

$$w_{\text{tot}} = \sum_{(s,w_s)\in W} w_s = \sum_{(s',w'_{s'})\in W'} w'_{s'}.$$

The two weighted sets W and W' define two distributions $\pi_{S,W}$ and $\pi_{S,W'}$ on S. We want to specify a coupling Ψ of $\pi_{S,W}$ and $\pi_{S,W'}$. Let $K_{|S|,|S|}$ be an edge-weighted complete bipartite graph with vertex sets W and W'. That is, for each pair $(s, w_s) \in W$ there is an edge to every pair in W'. Every edge e of $K_{|S|,|S|}$ has a weight $w_e \geq 0$ such that the following condition holds. Let (s, w_s) be any vertex in $K_{|S|,|S|}$ and let E be the set of all $|S|$ edges incident to (s, w_s). Then $\sum_{e\in E} w_e = w_s$. It follows that the sum of the edge weights of all $|S|^2$ edges in $K_{|S|,|S|}$ equals w_{tot}. The bipartite graph $K_{|S|,|S|}$ represents a coupling Ψ of $\pi_{S,W}$ and $\pi_{S,W'}$. In order to draw a pair of elements from Ψ, we randomly select an edge e in $K_{|S|,|S|}$ with probability w_e/w_{tot}. If edge $e = ((s, w_s), (s', w'_{s'}))$ is chosen, it means that we have drawn s from $\pi_{S,W}$ and s' from $\pi_{S,W'}$.

10.3.2 The proof of Lemma 10.2

The bipartite graph representation of a coupling will be used when we construct couplings of colourings of 2×2-blocks in the proof of Lemma 10.2. We restate the lemma before giving the proof.

Lemma (10.2). *Let $v_1, \ldots, v_4$ be the four vertices in a 2×2-block and let $z_1, \ldots, z_8$ be the boundary vertices of the block and let the labelling be as in Figure 10.1. Let Z and Z' be any two 7-colourings of the boundary vertices such that Z and Z' are identical on each vertex except on z_1. Let π_Z and $\pi_{Z'}$ be the uniform distributions on proper 7-colourings of the block that agree with Z and Z', respectively. For $i = 1, \ldots, 4$ let $p_{v_i}(\Psi)$ denote the probability that the colour of vertex v_i differ in a pair of colourings drawn from a coupling Ψ of π_Z and $\pi_{Z'}$. Then there exists a coupling Ψ such that $p_{v_1}(\Psi) < 0.283$, $p_{v_2}(\Psi) < 0.079$, $p_{v_3}(\Psi) < 0.051$ and $p_{v_4}(\Psi) < 0.079$.*

Proof. Fix two colourings Z and Z' of the boundary that differ only on vertex z_1. Let c be the colour of vertex z_1 in Z and let $c' \neq c$ be the colour of z_1 in Z'. Let $\Omega(Z)$ and $\Omega(Z')$ be the two sets of proper 7-colourings of the block that agree with Z and Z', respectively. For ease of notation in this proof, let Ω^+ denote the set of all 7-colourings of the block. Let W_Z and $W_{Z'}$ be two weighted sets of Ω^+. The weights are assigned as follows.

- For the pair $(\sigma, w_\sigma) \in W_Z$ let the weight $w_\sigma = |\Omega(Z')|$ if $\sigma \in \Omega(Z)$, otherwise let $w_\sigma = 0$.
- For the pair $(\sigma, w_\sigma) \in W_{Z'}$ let the weight $w_\sigma = |\Omega(Z)|$ if $\sigma \in \Omega(Z')$, otherwise let $w_\sigma = 0$.

It follows from the assignment of the weights that the distribution π_{Ω^+, W_Z} is the uniform distribution on $\Omega(Z)$. That is, $\pi_{\Omega^+, W_Z} = \pi_Z$. Similarly, $\pi_{\Omega^+, W_{Z'}}$ is the uniform distribution $\pi_{Z'}$ on $\Omega(Z')$. Note that $w_{\text{tot}} = |\Omega(Z)||\Omega(Z')|$. Then a coupling Ψ of π_{Ω^+, W_Z} and $\pi_{\Omega^+, W_{Z'}}$ is specified with an edge-weighted complete bipartite graph $K = K_{|\Omega^+|, |\Omega^+|}$ as described in the previous section. Let E be the set of all edges $e = ((\sigma, w_\sigma), (\sigma', w'_{\sigma'}))$ in K such that σ and σ' differ only on vertex v_i. Then $p_{v_i}(\Psi) = \sum_{e \in E} w_e / w_{\text{tot}}$.

In order to obtain sufficiently small upper bounds on $p_{v_i}(\Psi)$ for the four vertices $v_1, \ldots, v_4$ in the block, we would like to assign weights to the edges of K such that as much weight as possible is assigned to edges between colourings that agree on many vertices in the block. In general it is not clear exactly how to do this. For instance, if we assign too much weight on edges between colourings that are identical on vertex v_2 then we might not be able to assign as much weight as we would like to on edges between colourings that are identical on vertex v_4. Thus, the probability of having a mismatch on v_4 would increase. Intuitively, a good strategy would be to assign as much weight as possible on edges between colourings that are identical on the whole block. This implies that we try to assign as much weight as possible on edges between colourings that are identical on vertex v_1, the vertex adjacent to the discrepancy vertex z_1 on the boundary. If there is a mismatch on vertex v_1 it should be a good idea to assign

as much weight as possible on edges between colourings that are identical on the whole block apart from vertex v_1. This idea leads to a heuristic in which the assignment of the edge weights is divided into three phases. The exact procedure is described next.

In phase one we match identical colourings. For all colourings $\sigma \in \Omega^+$ of the block the edge $e = ((\sigma, w_\sigma), (\sigma, w'_\sigma))$ in K will be given weight $w_e = \min(w_\sigma, w'_\sigma)$. That is, we maximise the probability of choosing the same colouring σ from both π_{Ω^+, W_Z} and $\pi_{\Omega^+, W_{Z'}}$.

For the following two phases we define an ordering of the colourings in Ω^+. We order the colourings lexicographically with respect to the vertex order v_3, v_2, v_4, v_1. That is, if the seven colours are $1, \ldots, 7$ then the colouring of v_3, v_2, v_4, v_1 will start with 1, 1, 1, 1, respectively. The next colouring will be 1, 1, 1, 2, and so on. This ordering of colourings in Ω^+ carries over to an ordering of the pairs in W_Z and $W_{Z'}$. That is, we order the pairs (σ, w_σ) in W_Z with respect to the lexicographical ordering of σ. Similarly we order the pairs in $W_{Z'}$. This ordering of the pairs will be important in the next two phases. It provides some level of control of how colourings are being paired up in terms of the assignment of weights on edges in K. Edges will be considered with respect to this ordering because an arbitrary ordering of the edges would not necessarily result in probabilities $p_{v_i}(\Psi)$ that would be within the bounds of the lemma.

In the second phase we ignore the colour of vertex v_1 and match colourings that are identical on all of the remaining three vertices v_2, v_3 and v_4. More precisely, for each pair $(\sigma, w_\sigma) \in W_Z$, considered in the order explained above, we consider the edges $e = ((\sigma, w_\sigma), (\sigma', w'_{\sigma'}))$ where σ and σ' are identical on all vertices but v_1. The edges are considered according to the ordering of the second component $(\sigma', w'_{\sigma'}) \in W_{Z'}$, keeping the first component fixed. We assign as much weight as possible to e such that the total weight on edges incident to $(\sigma, w_\sigma) \in W_Z$ does not exceed w_σ and such that the total weight on edges incident to $(\sigma', w'_{\sigma'}) \in W_{Z'}$ does not exceed $w'_{\sigma'}$. Note that in the lexicographical ordering of the colourings, vertex v_1 is the least significant vertex and therefore the ordering provide some level of control of how to pair up colourings that are similar on the remaining three vertices.

In the third and last phase we assign the remaining weights on the edges. As in phase two, for each pair $(\sigma, w_\sigma) \in W_Z$ we consider the edges $e = ((\sigma, w_\sigma), (\sigma', w'_{\sigma'}))$. The pairs and edges are considered in accordance with the ordering explained above. The difference between the second and third phase is that now we do not have any restrictions on the colourings σ and σ'. We assign as much weight as possible to e such that the total weight on edges incident to $(\sigma, w_\sigma) \in W_Z$ does not exceed w_σ and such that the total weight on edges incident to $(\sigma', w'_{\sigma'}) \in W_{Z'}$ does not exceed $w'_{\sigma'}$. After phase three we have assigned all weights to the edges of K and hence K represents a coupling Ψ of π_Z and $\pi_{Z'}$. It turns out that the resulting coupling is sufficiently good

for proving the lemma.

From K we compute the probabilities $p_{v_1}(\Psi)$, $p_{v_2}(\Psi)$, $p_{v_3}(\Psi)$ and $p_{v_4}(\Psi)$ by summing edge weights as described above. We have written a C-program which loops through all colourings Z and Z' of the boundary of the block and constructs the bipartite graph K as described here. For each boundary the probabilities $p_{v_1}(\Psi)$, $p_{v_2}(\Psi)$, $p_{v_3}(\Psi)$ and $p_{v_4}(\Psi)$ are successfully verified to be within the bounds of the lemma. Details on the C-program are referred to the webpage `http://www.csc.liv.ac.uk/~markus/phdthesis/` □

10.4 Partial results for 6-colourings of the grid

In previous sections we have seen that systematic scan on the grid using 2×2-blocks and seven colours mixes rapidly. An immediate question is whether we can do better and show rapid mixing with six colours. This matter will be discussed next, and we will show that even with bigger block sizes (up to 3×3), it is not possible to show rapid mixing using the technique described in this chapter. More precisely, we will establish lower bounds on the parameter α for 2×2-blocks, 2×3-blocks and 3×3-blocks. All three lower bounds are greater than one and hence we cannot make use of Theorem 10.1 to show rapid mixing. We would need to either consider larger blocks or use a different approach. Considering larger blocks is not a feasible option since the number of required computations would result in a computer program with too long running time.

10.4.1 Establishing lower bounds for 2×2 blocks

We start by examining the 2×2-block again but this time with six colours. Lemma 10.2 provides upper bounds (under any colourings of the boundary) on the probabilities of having discrepancies at each of the four vertices of the block when two 7-colourings are drawn from the specified coupling. For six colours we will show lower bounds on these probabilities under any coupling and a specified pair of boundary colourings. Once again, let $v_1, \ldots, v_4$ be the four vertices in a 2×2-block and let $z_1, \ldots, z_8$ be the boundary vertices of the block, and let the labelling be as in Figure 10.1. Let Z and Z' be any two 6-colourings of the boundary vertices that assign the same colour to each vertex, except for z_1 on which they differ. Let π_Z and $\pi_{Z'}$ be the uniform distributions on the sets of proper 6-colourings of the block that agree with Z and Z', respectively. For $k = 1, \ldots, 4$, let $\Psi_{v_k}^{\min}(Z, Z')$ be a coupling of π_Z and $\pi_{Z'}$ that minimises $p_{v_k}(\Psi)$ over all couplings Ψ of π_Z and $\pi_{Z'}$. Let $p_{v_k}^{\text{low}} = \max_{Z,Z'} p_{v_k}(\Psi_{v_k}^{\min}(Z, Z'))$, where the maximum is taken over all pairs (Z, Z') of colourings of the boundary that differ only on z_1. From the definition of $p_{v_k}^{\text{low}}$ it follows that, for each v_k, there are two 6-colourings

Z and Z' of the boundary of the 2×2-block that differ only on z_1 such that, for all couplings Ψ of π_Z and $\pi_{Z'}$ we have $p_{v_k}(\Psi) \geq p_{v_i}^{\text{low}}$. We have the following lemma, which is proved with computer assistance.

Lemma 10.4. *Consider 6-colourings of the 2×2-block in Figure 10.1. We have* $p_{v_1}^{\text{low}} \geq 0.379$, $p_{v_2}^{\text{low}} \geq 0.107$, $p_{v_3}^{\text{low}} \geq 0.050$ *and* $p_{v_4}^{\text{low}} \geq 0.107$.

Proof. Fix one vertex v_k in the block and fix two 6-colourings Z and Z' of the boundary of the block such that they differ only on z_1. Let $\Omega(Z)$ and $\Omega(Z')$ be the two sets of proper 6-colourings of the block that agree with Z and Z', respectively. Let π_Z and $\pi_{Z'}$ be the uniform distributions on $\Omega(Z)$ and $\Omega(Z')$, respectively. For $c = 1, \ldots, 6$, let w_c be the number of colourings in $\Omega(Z)$ for which vertex v_k is assigned colour c. Similarly, let w'_c be the number of colourings in $\Omega(Z')$ for which vertex v_k is assigned colour c. Let $\pi_{Z,v_k}(c)$ be the probability that vertex v_k receives colour c in a colouring drawn from π_Z. Then $\pi_{Z,v_k}(c) = w_c/|\Omega(Z)|$. For $c = 1, \ldots, 6$, let $m_c = w_c|\Omega(Z')|$, $m'_c = w'_c|\Omega(Z)|$ and $M = |\Omega(Z)||\Omega(Z')|$. It follows that $\pi_{Z,v_k}(c) = m_c/M$ and $\pi_{Z',v_k}(c) = m'_c/M$. Observe that the quantities m_c, m'_c and M can be easily computed for a given pair of boundary colourings.

Now, let Ψ be any coupling of π_Z and $\pi_{Z'}$. The probability that vertex v_k is coloured c in both colourings drawn from Ψ can be at most $\min(m_c, m'_c)/M$. Therefore, the probability of drawing two colourings from Ψ such that the colour of vertex v_k is the same in both colourings is at most $\sum_{c=1}^{6} \min(m_c, m'_c)/M$, and the probability of assigning different colours to vertex v_k is at least

$$p_{v_k}(\Psi) \geq 1 - \sum_{c=1}^{6} \frac{\min(m_c, m'_c)}{M}.$$

We have successfully verified the bounds in the statement of the lemma by maximising this lower bound on $p_{v_k}(\Psi)$ over all boundary colourings Z and Z' and for each vertex v_k in the block. The computations are carried out with the help of a computer program written in C. For details on the C-program, see webpage `http://www.csc.liv.ac.uk/~markus/phdthesis/` □

For seven colours, Corollary 10.3 makes use of Lemma 10.2 to establish upper bounds on the influence parameters $\rho_{u,v}^k$. These parameters are used in the proof of Theorem 3.6 to obtain an upper bound on the parameter α. The upper bound on α is shown to be less than one which implies rapid mixing for seven colours when applying Theorem 10.1. We can use Lemma 10.4 to obtain lower bounds on the influence parameters $\rho_{u,v}^k$ by completing the coupling in a way analogous to the coupling in Corollary 10.3. This in turn will result in a lower bound on the parameter α that is greater than one. That is, following the proof of Theorem 3.6 and making use of

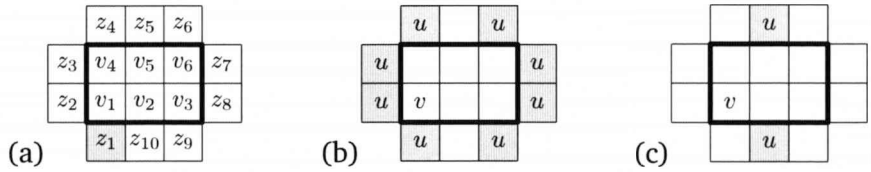


Figure 10.3. (a) General labelling of the vertices in a 2×3-block Θ_k and the vertices $\partial\Theta_k$ on the boundary of the block. (b)–(c) All ten positions of a vertex $u \in \partial\Theta_k$ on the boundary of the block in relation to a vertex $v \in \Theta_k$ in the corner of the block.

Lemma 10.4, a lower bound on α will be

$$\alpha \geq 2(0.379 + 0.107 + 0.050 + 0.107) = 1.286 > 1.$$

Hence we fail to show rapid mixing of systematic scan with six colours using 2×2-blocks.

10.4.2 Bigger blocks

We fail to show rapid mixing of systematic scan with six colours and 2×2-blocks, and we will now show that increasing the block size to both 2×3 and 3×3 will not be sufficient either. Lemma 10.5 below considers 2×3-blocks and is analogous to Lemma 10.4. We make use of the same notation as for Lemma 10.4, only the block is bigger and the labelling of the vertices is different (see Figure 10.3(a)). Lemma 10.5 is proved by computation in the same way as Lemma 10.4. For details on the C-program used in the proof, see webpage `http://www.csc.liv.ac.uk/~markus/phdthesis/`

Lemma 10.5. *Consider 6-colourings of the 2×3-block in Figure 10.3(a). We have* $p_{v_1}^{\text{low}} \geq 0.3671$, $p_{v_3}^{\text{low}} \geq 0.0298$, $p_{v_4}^{\text{low}} \geq 0.0997$ *and* $p_{v_6}^{\text{low}} \geq 0.0174$.

We will now use Lemma 10.5 to show that $\alpha > 1$ for 2×3 blocks. Let Θ_k be any 2×3-block and let $v \in \Theta_k$ be a vertex in a corner of the block. A vertex $u \in \partial\Theta_k$ on the boundary of the block can occupy ten different positions on the boundary in relation to v. See Figure 10.3(b) and (c). We determine lower bounds on the influences $\rho_{u,v}^k$ of u on v under Θ_k from Lemma 10.5. However, Lemma 10.5 provides lower bounds on $\rho_{u,v}^k$ only when $u \in \partial\Theta_k$ is adjacent to a corner vertex of the block, as in Figure 10.3(b). If u is located as in Figure 10.3(c) we do not know more than that $\rho_{u,v}^k$ is bounded from below by zero. Nevertheless, the lower bound on α exceeds one. Let $\alpha_{k,v} = \sum_{u \in \partial\Theta_k} \rho_{u,v}^k$ be the influence on v under Θ_k. Following the proof of Theorem 3.6

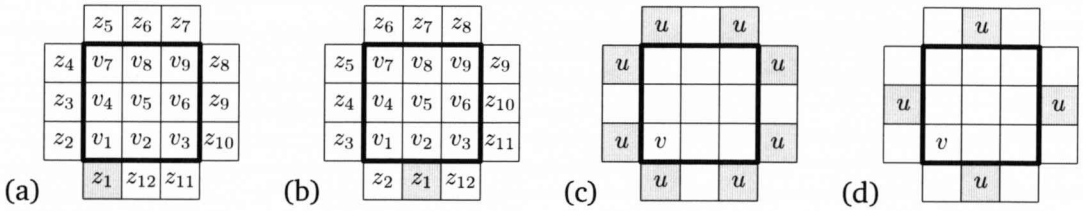


Figure 10.4. (a)–(b) General labelling of the vertices in a 3×3-block Θ_k and two different labellings of the vertices $\partial\Theta_k$ on the boundary of the block. The discrepancy vertex on the boundary has label z_1. (c)–(d) All twelve positions of a vertex $u \in \partial\Theta_k$ on the boundary of the block in relation to a vertex $v \in \Theta_k$ in the corner of the block.

and using the lower bounds in Lemma 10.5 we have

$$\begin{aligned}\alpha_{k,v} &= \sum_{u \text{ in Fig. 10.3(b)}} \rho^k_{u,v} + \sum_{u \text{ in Fig. 10.3(c)}} \rho^k_{u,v} \;\geq\; 2(0.3671 + 0.0298 + 0.0997 + 0.0174) \\ &= 1.028,\end{aligned}$$

where we set the lower bound on the second sum to zero. Thus, $\alpha \geq 1.028 > 1$ which means that we cannot use Theorem 10.1 to show rapid mixing of systematic scan with six colours and 2×3-blocks.

Lastly, we increase the block size to 3×3 and show that a lower bound on α is still greater than one. We have the following lemma which is proved by computation in the same way as Lemmas 10.4 and 10.5. For details on the C-program used in the proof, see webpage `http://www.csc.liv.ac.uk/~markus/phdthesis/`

Lemma 10.6. *For 6-colourings of the 3×3-block with vertices labelled as in Figure 10.4(a), we have* $p^{\text{low}}_{v_1} \geq 0.3537$, $p^{\text{low}}_{v_3} \geq 0.0245$, $p^{\text{low}}_{v_7} \geq 0.0245$ *and* $p^{\text{low}}_{v_9} \geq 0.0071$. *Furthermore, for 6-colourings of the 3×3-block in Figure 10.4(b), we have* $p^{\text{low}}_{v_1} \geq 0.0838$, $p^{\text{low}}_{v_3} \geq 0.0838$, $p^{\text{low}}_{v_7} \geq 0.0138$ *and* $p^{\text{low}}_{v_9} \geq 0.0138$.

Note that Lemma 10.6 provides lower bounds on the probabilities of having a mismatch on a corner vertex of the block when the discrepancy vertex on the boundary (labelled z_1) is adjacent to a corner vertex (Figure 10.4(a)) and adjacent to a middle vertex (Figure 10.4(b)). Let Θ_k be any 3×3-block and let $v \in \Theta_k$ be a vertex in a corner of the block. A vertex $u \in \partial\Theta_k$ on the boundary of the block can occupy twelve different positions in relation to v. See Figure 10.4(c) and (d). Analogous to Corollary 10.3, lower bounds on the influences $\rho^k_{u,v}$ of u on v under Θ_k can be determined from Lemma 10.6. Let $\alpha_{k,v} = \sum_u \rho^k_{u,v}$ be the influence on v under Θ_k. Following the

proof of Theorem 3.6 and using the lower bounds in Lemma 10.6, we have

$$\begin{aligned}\alpha_{k,v} &= \sum_{u \text{ in Fig. 10.4(c)}} \rho_{u,v}^{k} + \sum_{u \text{ in Fig. 10.4(d)}} \rho_{u,v}^{k} \\ &\geq 2(0.3537 + 0.0245 + 0.0245 + 0.0071) + (0.0838 + 0.0838 + 0.0138 + 0.0138) \\ &= 1.0148.\end{aligned}$$

Thus, $\alpha \geq 1.0148 > 1$, which means that we cannot use Theorem 10.1 to show rapid mixing of systematic scan with six colours and 3×3-blocks.

A natural question is whether we can show rapid mixing using even bigger blocks. It seems possible to do this although the computations rapidly become intractable as the block size increases. Already with a 3×3-block the number of boundary colourings we need to consider (after removing isomorphisms) is in excess of 10^6, and for each boundary colouring there are more than 10^7 colourings of the block to consider. In addition to the number of colourings to consider, the time it would take to construct the required couplings, as we did in the proof of Lemma 10.2, would increase. Also, when using a larger block size, different positions of the vertex v in the block need to be considered whereas we could nicely make use of the symmetry of the 2×2-block and only consider one position of v (a corner vertex). If different positions of v have to be considered this has to be captured in the construction of the coupling and would likely require more computations. The conclusion is that in order to show rapid mixing for six colours of systematic scan on the grid we would most likely have to rely on a different approach than the one presented in this chapter.

Chapter 11

Conclusion

In this thesis we give improved mixing bounds for spin systems under the anti-ferromagnetic Pott's model. The results are stated in Chapter 3. As pointed out in Chapter 3, there is room for improvement. Often the believed mixing bounds do not match the mixing bounds that haven been rigorously proved. For instance, Salas and Sokal [SS97] say that it is believed that strong spatial mixing occurs for the square lattice for every $q \geq 4$ and $\lambda \in [0, 1]$. To prove mixing for $q = 4$ and $q = 5$ for values of λ all the down to 0 therefore remains an open problem. For the triangular lattice is it believed that there is strong spatial mixing for every $q \geq 5$ and $\lambda \in [0, 1]$. As we have seen in this thesis, mixing has only been proved for $q \geq 9$. Furthermore, the proof of mixing for $q = 9$ and the proof of mixing for $q = 10$ [GMP05] only hold for $\lambda = 0$. For the kagome lattice it is believed that there is strong spatial mixing for every $q \geq 4$ and $\lambda \in [0, 1]$. That is, to prove mixing for $q = 4$ and small values of λ ($\lambda = 0$ in particular) is still an open problem.

The parameter λ controls the level of interaction between neighbouring vertices. For $\lambda = 0$ only proper colourings are allowed. It is reasonable to believe that if strong spatial mixing has been shown for $\lambda = \lambda_0$, where $\lambda_0 \in [0, 1]$, then there should be strong spatial mixing for every $\lambda \in [\lambda_0, 1]$. However, proving this is still an open problem. Recall from Chapter 7 how we proved mixing for an interval of λ by analysing small subintervals. The proof of Goldberg, Martin and Paterson's general theorem on strong spatial mixing for $\lambda = 0$ (Theorem 3.8) does not extend naturally to positive values of λ. It would be interesting to have more mixing results that span over a larger range of λ.

11.1 Further improvements

In this thesis we have shown mixing by defining suitable sets of edge-boundary pairs X and then using these sets with the recursive coupling tree to show that the quan-

tity $\Gamma^d(X)$ decreases exponentially with d. As explained in Section 6.3, the proof of Theorem 3.8 uses this technique with sets of edge-boundary pairs X whose region R_X only contains the vertex v_X and neighbours of v_X. From Theorem 3.8 we have strong spatial mixing for $\lambda = 0$ and the following values of q and Δ. The table below lists (approximate) lower bounds on q for $\Delta = 4, \ldots, 23$.

Δ	q	Δ	q	Δ	q	Δ	q
4	6.5826	9	15.3987	14	24.2148	**19**	**33.0309**
5	8.3458	10	17.1619	15	25.9780	20	34.7941
6	**10.1090**	11	18.9251	16	27.7412	21	36.5573
7	11.8723	12	20.6883	17	29.5044	22	38.3205
8	13.6355	13	22.4516	18	31.2677	**23**	**40.0838**

From the table above it follows that there is mixing for any 6-regular (triangle-free) graph with $q \geq 11$ colours. However, the value of q in the table for $\Delta = 6$ is only slightly above 10. This suggests that a slight improvement of the the general mixing result in Theorem 3.8 would push the mixing bound on q below 10 and thereby imply mixing for $q = 10$ colours and $\Delta = 6$. Goldberg, Martin and Paterson [GMP05] made this observation and used the recursive couple tree with a collection of only a handful sets of edge-boundary pairs to show mixing for the triangular lattice and the integer lattice $\mathbb{Z}^3$ for $q = 10$. These lattices are both 6-regular graphs. The small number of sets of edge-boundary pairs that were needed for this improvement did not require computer assistance when deriving the inequalities. However, computer assistance was used to compute some of the μ-values.

In the table above we have highlighted values when Δ is 6, 19 and 23. A small improvement of the general mixing result in Theorem 3.8 might imply that the mixing bound on q for $\Delta = 6$, $\Delta = 19$ and $\Delta = 23$ will drop from 11 to 10, 34 to 33 and from 41 to 40, respectively. An improvement of Theorem 3.8 could be possible by using sets of edge-boundary pairs X whose region R_X contains neighbours of neighbours of v_X. It is not obvious how to analyse this in a general setting with an arbitrary maximum degree Δ. How to maximise the μ-values for edge-boundary pairs X whose region R_X contains v_X, neighbours of v_X and neighbours of neighbours of v_X is a problem that must be dealt with. For small values of Δ it could be possible to do this with computer assistance, however, the number of edge-boundary pairs to consider would be rather large for $\Delta = 6$ already.

11.1.1 Some lattices seem trickier to work with than others

Theorem 3.8 assumes that the graph $\mathcal{G}$ is triangle-free. The reason for this assumption is that the absence of triangles made it easier to compute upper bounds on the μ-values

(see Section 6.3). An interesting observation is that the presence of triangles seems to make it easier to derive recurrences from the recursive coupling tree.

Let X be an edge-boundary pair. When branching in the recursive coupling tree T_X we recursively consider neighbours of v_X. A path in T_X from the root corresponds to a self-avoiding walk in R_X, where a self-avoiding walk is a path from one vertex to another which never intersects itself. The presence of many triangles in the graph clearly reduces the number of short self-avoiding walks. Both the triangular lattice and the kagome lattice contain triangles. These triangles played an important role in the mixing proofs in Chapter 8 and Chapter 9 since they kept the branching factor in the tree T_X sufficiently low. A low branching factor naturally allows higher μ-values. This has been discussed extensively in Chapter 6.

On the contrary to the triangular lattice and the kagome lattice, the square lattice does not contain triangles. The shortest path in the square lattice that intersects itself has length 4. The kagome lattice and the square lattice both have maximum degree $\Delta = 4$. However, the absence of triangles in the square lattice implies a higher branching factor in the recursive coupling tree compared to the kagome lattice. A higher branching factor requires lower μ-values. Lower μ-values are obtained by considering larger regions. Larger regions imply more computation. This is probably the reason why we have failed to prove strong spatial mixing for $q = 5$ and $\lambda = 0$ on the square lattice but succeeded in doing so for the kagome lattice. We have tried to define good $(\mathcal{A}, \mathcal{F})$-sets by considering large regions of the square lattice. The required computations have turned out to be too involved to be feasible to work with. The idea of incorporating colours of boundary edges in the recursions (as for the kagome lattice in Chapter 9) might be a possible approach. However, the lack of triangles does not allow for the same straightforward use of edge colours. Again, the required computations do not seem feasible to deal with in practice.

We have tried and failed to show strong spatial mixing for $q = 9$ and $\lambda = 0$ on the lattice $\mathbb{Z}^3$. As for the triangular lattice, $\mathbb{Z}^3$ is 6-regular, however, the lack of triangles implies a higher branching factor. Another disadvantage with $\mathbb{Z}^3$ is that it is not planar. Planar graphs allow for a straightforward visualisation of regions. If the graph is not planar then it is harder to visually get a good intuitive feeling for the recursions in the recursive coupling tree, especially when the number of sets of edge-boundary pairs that are considered is large. When Goldberg, Martin and Paterson [GMP05] showed strong spatial mixing for $q = 10$ and $\lambda = 0$ on $\mathbb{Z}^3$ they worked with a small collection of sets of edge-boundary pairs.

If one attempts to improve the mixing bounds for a specific lattice using the techniques in this thesis then it is worth to keep in mind that some lattices are more "friendly" to work with than others; lattices with triangles seem to be easier than lattices without triangles, and planar graphs allow for a straightforward visualisation that

might be important when the sizes of the regions grow and the complexity of the recurrences increases.

Bibliography

[Ald83] D. Aldous. Random walks on finite groups and rapidly mixing Markov chains. In *Seminar on Probability, XVII*, volume 986 of *Lecture Notes in Mathematics*, pages 243–297. Springer, 1983.

[AMMV04] D. Achlioptas, M. Molloy, C. Moore, and F. Van Bussel. Sampling grid colorings with fewer colours. In *LATIN 2004: Theoretical Informatics*, volume 2976 of *Lecture Notes in Computer Science*, pages 80–89. Springer, 2004.

[BD97a] R. Bubley and M. Dyer. Path coupling: a technique for proving rapid mixing in Markov chains. In *FOCS '97: Proceedings of the 38th Symposium on Foundations of Computer Science*, pages 223–231. IEEE Computer Society Press, 1997.

[BD97b] R. Bubley and M. Dyer. Path coupling, Dobrushin uniqueness, and approximate counting. Technical Report 97.04, School of Computing, University of Leeds, Leeds, United Kingdom, 1997.

[BDGJ99] R. Bubley, M. Dyer, C. Greenhill, and M. Jerrum. On approximately counting colourings of small degree graphs. *SIAM Journal on Computing*, 29(2):387–400, 1999.

[BW99] G. R. Brightwell and P. Winkler. Graph homomorphisms and phase transitions. *Journal of Combinatorial Theory, Series B*, 77(2):221–262, 1999.

[Ces01] F. Cesi. Quasi-factorization of the entropy and logarithmic Sobolev inequalities for Gibbs random fields. *Probability Theory and Related Fields*, 120:569–584, 2001.

[DG99] M. Dyer and C. Greenhill. Random walks on combinatorial objects. In *Surveys in Combinatorics*, volume 267 of *London Mathematical Society Lecture Note Series*, pages 101–136. Cambridge University Press, 1999.

[DG00] M. Dyer and C. Greenhill. On Markov chains for independent sets. *Journal of Algorithms*, 35(1):17–49, 2000.

[DGJ06a] M. Dyer, L. A. Goldberg, and M. Jerrum. Dobrushin conditions and systematic scan. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques*, volume 4110 of *Lecture Notes in Computer Science*, pages 327–338. Springer, 2006.

[DGJ06b] M. Dyer, L. A. Goldberg, and M. Jerrum. Systematic scan and sampling colourings. *Annals of Applied Probability*, 16(1):185–230, 2006.

[DGJM06] M. Dyer, L. A. Goldberg, M. Jerrum, and R. Martin. Markov chain comparison. *Probability Surveys*, 3:89–111, 2006.

[Dia88] P. Diaconis. *Group representations in probability and statistics*. Institute of Mathematical Statistics, Hayward, CA, USA, 1988.

[Dob68] R. L. Dobrushin. Gibbsian random fields for lattice systems with pairwise interactions. *Functional Analysis and Its Applications*, 2(4):292–301, 1968.

[DS91] P. Diaconis and D. Stroock. Geometric bounds for eigenvalues of Markov chains. *Annals of Applied Probability*, 1(1):36–61, 1991.

[DSC93] P. Diaconis and L. Saloff-Coste. Comparison theorems for reversible Markov chains. *Annals of Applied Probability*, 3(3):696–730, 1993.

[DSVW04] M. Dyer, A. Sinclair, E. Vigoda, and D. Weitz. Mixing in time and space for lattice spin systems: a combinatorial view. *Random Structures and Algorithms*, 24(4):461–479, 2004.

[Geo88] H.-O. Georgii. *Gibbs measures and phase transitions*. de Gruyter Studies in Mathematics 9. Walter de Gruyter & Co., Berlin, Germany, 1988.

[GHM01] H.-O. Georgii, O. Häggström, and C. Maes. The random geometry of equilibrium phases. *Phase Transitions and Critical Phenomena*, 18:1–142, 2001.

[GJ90] M. R. Garey and D. S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness*. W. H. Freeman & Co., New York, NY, USA, 1990.

[GJMP06] L. A. Goldberg, M. Jalsenius, R. Martin, and M. Paterson. Improved mixing bounds for the anti-ferromagnetic potts model on $\mathbb{Z}^2$. *LMS Journal of Computation and Mathematics*, 9:1–20, 2006.

[GMP04] L. A. Goldberg, R. Martin, and M. Paterson. Random sampling of 3-colourings in $\mathbb{Z}^2$. *Random Structures and Algorithms*, 24(3):279–302, 2004.

[GMP05] L. A. Goldberg, R. Martin, and M. Paterson. Strong spatial mixing with fewer colours for lattice graphs. *SIAM Journal on Computing*, 35(2):486–517, 2005.

[Isi25] E. Ising. Beitrag zur Theorie der Ferromagnetismus. *Zeitschrift für Physik*, 31:253–258, 1925.

[Jer95] M. Jerrum. A very simple algorithm for estimating the number of k-colorings of a low-degree graph. *Random Structures and Algorithms*, 7(2):157–165, 1995.

[Jer03] M. Jerrum. *Counting, Sampling and Integrating: Algorithms and Complexity*. Birkhäuser, Basel, Switzerland, 2003.

[JVV86] M. Jerrum, L. Valiant, and V. Vazirani. Random generation of combinatorial structures from a uniform distribution. *Theoretical Computer Science*, 43:169–188, 1986.

[Len20] W. Lenz. Beitrag zum Verständnis der magnetischen Erscheinungen in festen Körpern. *Physikalische Zeitschrift*, 21:613–615, 1920.

[LR69] O. Lanford and D. Ruelle. Observables at infinity and states with short range correlations in statistical mechanics. *Communications in Mathematical Physics*, 13(3):194–215, 1969.

[LRS01] M. Luby, D. Randall, and A. Sinclair. Markov chain algorithms for planar lattice structures. *SIAM Journal on Computing*, 31(1):167–192, 2001.

[Mar99] F. Martinelli. Lectures on Glauber dynamics for discrete spin models. In *Lectures on Probability Theory and Statistics (Saint-Flour, 1997)*, volume 1717 of *Lecture Notes in Mathematics*, pages 93–191. Springer, 1999.

[MO94a] F. Martinelli and E. Olivieri. Approach to equilibrium of Glauber dynamics in the one phase region I: The attractive case. *Communications in Mathematical Physics*, 161(3):447–486, 1994.

[MO94b] F. Martinelli and E. Olivieri. Approach to equilibrium of Glauber dynamics in the one phase region II: The general case. *Communications in Mathematical Physics*, 161(3):487–514, 1994.

[Pap94] C. M. Papadimitriou. *Computational complexity*. Addison-Wesley, Reading, MA, USA, 1994.

[Ped07] K. Pedersen. Dobrushin conditions for systematic scan with block dynamics. In *Mathematical Foundations of Computer Science 2007*, volume 4708 of *Lecture Notes in Computer Science*, pages 264–275. Springer, 2007.

[Pot52] R. B. Potts. Some generalized order-disorder transformations. *Proceedings of the Cambridge Philosophy Society*, 48:106–109, 1952.

[Ros06] S. M. Ross. *Introduction to Probability Models, Ninth Edition*. Academic Press, Inc., Orlando, FL, USA, 2006.

[RT00] D. Randall and P. Tetali. Analyzing Glauber dynamics by comparison of Markov chains. *Journal of Mathematical Physics*, 41:1598–1615, 2000.

[Sin92] A. Sinclair. Improved bounds for mixing rates of Markov chains and multicommodity flow. *Combinatorics, Probability and Computing*, 1:351–370, 1992.

[SS97] J. Salas and A. D. Sokal. Absence of phase transition for antiferromagnetic Potts models via the Dobrushin uniqueness theorem. *Journal of Statistical Physics*, 86(3–4):551, 1997.

[SZ92] D. W. Stroock and B. Zegarlinski. The logarithmic Sobolev inequality for discrete spin systems on a lattice. *Communications in Mathematical Physics*, 149(1):175–193, 1992.

[Val79] L. G. Valiant. The complexity of computing the permanent. *Theoretical Computer Science*, 8:189–201, 1979.

[Vig01] E. Vigoda. A note on the Glauber dynamics for sampling independent sets. *Electronic Journal of Combinatorics*, 8(1), 2001.

[Wei04] D. Weitz. *Mixing in Time and Space for Discrete Spin Systems*. PhD thesis, University of California, Berkley, 2004.

[Wei05] D. Weitz. Combinatorial criteria for uniqueness of Gibbs measures. *Random Structures and Algorithms*, 27(4):445–475, 2005.

[Wei06] D. Weitz. Counting independent sets up to the tree threshold. In *STOC '06: Proceedings of the 38th Annual ACM Symposium on Theory of Computing*, pages 140–149. ACM Press, 2006.